

Engineered viral vectors and developed tissue  
clearing methods for single-cell phenotyping in  
whole organs

Thesis by  
Ken Yee Chan

In Partial Fulfillment of the Requirements for  
the degree of  
Doctor of Philosophy

The Caltech logo, featuring the word "Caltech" in a bold, orange, sans-serif font.

CALIFORNIA INSTITUTE OF TECHNOLOGY  
Pasadena, California

2017  
Defended May 18<sup>th</sup>, 2017

© 2017

Ken Yee Chan  
ORCID: 0000-0002-8853-5186

## ACKNOWLEDGMENTS

I begin by thanking my Ph.D. advisor, Dr. Viviana Gradinaru, for her guidance, hard work, dedication, persistence, and patience over the course of my Ph.D. training. Throughout my training, Viviana has advised me a great deal both professionally and personally. Viviana sets the bar high for each of her mentees, but because of her hard work, calm and caring nature, and dedication, she has provided me with the necessary means to achieve my goals during my Ph.D. training. She has done this by providing invaluable guidance and advice, an uncanny amount of resources, and a positive lab environment for learning, individual growth, and collaboration. Her patience and guidance has allowed me to learn many of the most up-to-date techniques in neuroscience, as well as valuable off-the-bench skills, such as writing papers, grants, and abstracts for conferences. She has also nurtured my future career by providing me with opportunities in presenting my work at numerous conferences and also allowing me to participate in many collaborations with other top scientists around the country. I am convinced my journey as a graduate student would not have been as productive and eventful had Viviana not taken the chance on accepting me into her lab. I hope her journey in mentoring me has been as positive and enjoyable as much as it has been for me in being her graduate student.

Furthermore, I would like to thank Dr. Benjamin E. Deverman. Ben has played a key role in my Ph.D. training and life in graduate school. His thoroughness, integrity, and efficiency on and off the bench have always been traits I have greatly admired. Ben has become one of the very few role models I find myself falling back to in times of self-reflection and growth. In addition, I have appreciated Ben's style of mentorship in my graduate studies. His mentorship has always come in the form of bringing others to the same level as his by training or advising a person as if he or she will eventually become his peer. Ben's cool headedness and unbelievable strive to become exceptional at whatever he sets out to do, in and outside of science, has inspired many facets of my life. I am extremely grateful to know Ben and to have worked with him but, more importantly, for not leaving me to my death during our bike rides up Glendora Mountain Road.

There are a few people without whom much of the work presented in this thesis would not have been possible, including my colleague and dear friend Dr. Alon Greenbaum. I have worked closely with Alon on many projects and have enjoyed and appreciated his candor and determination in getting things done. Alon's need for perfection, consistency, and efficiency has been a major backbone in our work together. I feel very fortunate to have worked alongside Alon on our projects. One of the most important things I have learned from Alon was how to carry out a project and finish it. I have always felt fortunate to have been paired with Alon and would never want to miss an opportunity to work alongside him, if possible.

I would also like to thank a few other postdoctoral scholars in Viviana's lab: Dr. Cheng Xiao, Dr. Jennifer Treweek, and Dr. Min Jee Jang. Cheng and Jennifer were both present when I first joined Viviana's lab. Much of my initial training in neuroscience techniques was largely from Cheng and Jennifer. Their investment in training me has played a large role in my success as a graduate student. I would also like to thank Min Jee Jang for her friendship, support, and help in my projects. She has been an amazingly helpful colleague and friend.

I would also like to thank Bin Yang. Bin was a research technician in Viviana's lab when I started graduate school. Bin had also taught me many things on the bench. Most importantly though, I am grateful that Bin had set up Viviana's lab so promptly. Surely, without a functional lab during my rotation, Viviana might not have taken me into her lab.

Graduate school has its ups and downs, and I believe I could not have dealt with this roller coaster experience without the many friends that I have met along the way. Two in particular come to mind. I would like to thank Nicholas Flytzanis and Ryan Cho for their friendship, humor, and support. Since both Nick and Ryan are also graduate students, they also understand the ups and downs of graduate school. Their role in being good friends has been invaluable. Their humor has been cheerful. Their support has been immensely helpful. I feel fortunate to have met them. I would also like to thank my friends and colleagues

Sripriya Ravindra Kumar, Bryan Yoo, Pradeep Rajendran, and Claire Bedbrook, who have always been there when I needed them. In addition, I would also like to thank the entire Gradinaru lab for being an amazing and well-rounded group of individuals that has supported me throughout my training.

In addition, I would like to thank all of the collaborators with whom I have worked with during my graduate training. In particular, the outstanding scientists Dr. Rogely Boyce and Dr. Helen McBride from Amgen for their, kindness, enthusiasm, valuable time, and generosity in sharing knowledge. I would also like to thank Dr. Carlos Lois, Dr. Victoria Orphan, Dr. Sarkis Mazmanian, and Dr. Luis Oscar Sanchez Guardado for collaborations in projects.

Furthermore, I would like to thank Dr. Henry Lester, Dr. Long Cai and Dr. John Allman for being on my oral candidacy and thesis committee. Henry, Long, and John have been guiding me, supporting my learning, and encouraging me over the last few years. I am extremely grateful to have such support during my Ph.D. training. In particular, I would like to thank Henry for his candor and enthusiasm over my work throughout the years and for always encouraging me to reach for my goals.

I would like to thank my land lord and lady, Martin Zitter and Carol Nakamoto-Zitter, whom I practically consider family now. I would also really like to thank them for not raising my rent in the past several years since I have been living with them.

In addition, I would like to thank my mom and dad for their continued support, and my friends Michael Knapp, Kristin Cummings, Tai Ting-Chi, Chris Tucker, Irina Meininger, Daniel Zitter, Dubravka Pezic, and Evelyn Stuwe. I would also like to thank my past mentors Dr. Casey Quinlan, Dr. Alexei Aravin, Dr. Bruce Hay, Dr. Katalin Fejes Tóth, Dr. Nikolai Kandul, and Dr. Ivan Olovnikov for their support, friendship, and guidance.

## ABSTRACT

A central question in biology is how different cell types interact with each other and their native environment to form complex functional systems and networks. Although our ability to investigate this question has considerably expanded from the development of genetically encoded tools, some limitations still persist. For instance, we are limited in our ability to visualize the native three dimensional environments of whole organs. Additionally, it is challenging to efficiently deliver transgenes into difficult-to-target areas through direct-injections, such as the cardiac ganglia, or broadly distributed networks, such as the myenteric nervous system, which limits our ability to extensively study these areas. Therefore, tools and methods that overcome these limitations are needed. Towards this end, my thesis work has been focused on developing tools for single-cell resolution phenotyping in whole organs. I have been developing tissue clearing technologies to render whole organs transparent for optical interrogation and characterizing viral capsids and engineering viral vectors for noninvasive widespread gene delivery to the central and peripheral nervous system.

Tissue clearing techniques for three dimensional optical interrogation were invented over a century ago. However, these earlier methods used harsh organic chemicals and failed to retain the tissue's native fluorescence or epitopes. These earlier methods eventually became incompatible to the hundreds of newly generated transgenic mouse lines that allowed for cell type-specific expression of fluorescent transgenes or to fluorescent labeling techniques, such as immunohistochemistry (IHC). The first part of my dissertation is aimed at addressing these limitations by further developing and standardizing a tissue clearing method that utilizes the vasculature to perfuse clearing reagents. This technique, called perfusion assisted agent release in situ (PARS) enables *(i)* whole organ clearing of soft tissue, *(ii)* preservation of native fluorescence, and *(iii)* preservation of epitopes compatible with IHC.

Although PARS allows us to optically investigate whole soft tissue organs, it is unsuitable for clearing bone tissue. The clearing of bone is important as it may provide

optical access to delicate environments, such as the lymphatic vessels lining the dural sinuses beneath the skull that would otherwise be damaged through traditional methods. However, clearing bone tissue is challenging since it is composed of both soft (bone marrow) and hard (mineral) tissue. To overcome this challenge, I developed a clearing method that rendered intact bone tissue transparent by using EDTA to decalcify bones and by constructing a convective flow chamber to efficiently clear bones. This method, called Bone CLARITY, is able to preserve native fluorescence and epitopes. In order to demonstrate the utility of Bone CLARITY, I collaborated with colleagues to quantitatively access a rare and non-uniformly distributed population of osteoprogenitor cells in their native three dimensional environment. Bone CLARITY in conjunction with light-sheet microscope enabled the early detection of an increase to this osteoprogenitor population after administration of a novel anabolic drug, which may have been undetected with traditional techniques.

Towards my second goal, I have been working on characterizing adeno-associated viruses (AAVs) for non-invasive widespread gene delivery across the central or peripheral nervous system. Through systemic delivery, these novel AAVs are able to efficiently deliver transgenes to (i) difficult-to-target areas, such as the dorsal root ganglia; (ii) cellular populations that are widely distributed across the mouse body, such as neurons in the myenteric nervous system, and (iii) through highly selective barriers, such as the blood-brain barrier. These viruses enable rapid expression of transgenes for perturbing and monitoring cellular circuits, or for potentially treating neurological diseases. In addition, I worked on engineering or validating several different gene regulatory elements to achieve cell type restricted expression in transgenic and non-transgenic animals with AAVs. These viral vectors may prove useful in rapidly testing newly developed genetic tools. Finally, I developed and characterized two different two-component viral vector systems to control the density of labeling when systemically delivering genes with our novel engineered viruses. I utilized this two-component system to perform single-cell morphology studies in the CNS and PNS. Collectively, these capsids and vectors expand the AAV toolbox and enable efficient and versatile gene delivery into the CNS and PNS of transgenic and non-transgenic animals.

## PUBLISHED CONTENT AND CONTRIBUTIONS

[1] Chan, K.Y.C et al. “Engineered adeno-associated viruses for efficient and noninvasive gene delivery throughout the central and peripheral nervous systems”. In press.

K.Y.C helped design and perform experiments, collect and analyze data, prepare figures and write the manuscript with input from all authors

[2] Greenbaum, A. et al. “Bone CLARITY: Clearing, imaging, and computational analysis of osteoprogenitors within intact bone marrow”. In: *Science Translational Medicine* 387.9 (2017). URL: <http://doi.org/10.1126/scitranslmed.aah6518>.

K.Y.C. performed experiments, data acquisition, data analysis, generated figures, and wrote the manuscript with input from all authors.

[3] Allen, W. et al. “Global Representations of Goal-Directed Behavior in Distinct Cell Types of Mouse Neocortex”. In: *Neuron* 94.4 (2017), pp. 891-907. URL: <http://dx.doi.org/10.1016/j.neuron.2017.04.017>.

K.Y.C. prepared AAV-PHP.B viruses.

[4] Xiao, C. et al. “Cholinergic Mesopontine Signals Govern Locomotion and Reward through Dissociable Midbrain Pathways”. In: *Neuron* 90.2 (2016), pp. 333-347. URL: <http://doi.org/10.1016/j.neuron.2016.03.028>

K.Y.C. contributed to animal behavior tests and data illustration.

[5] Deverman, B.E. et al. “Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain”. In: *Nature Biotechnology* 34.2 (2016), pp. 204-209. doi: 10.1038/nbt.3440.

K.Y.C. performed experiments, virus production and characterization.

[6] Treweek, J.B. et al. “Whole-body tissue stabilization and selective extractions via tissue-hydrogel hybrids for high-resolution intact circuit mapping and phenotyping”. In: *Nature Protocols* 10.11 (2015), pp. 1860-1896. doi:10.1038/nprot.2015.122.

K.Y.C designed and performed experiments, analyzed the data and prepared figures.

[7] Skennerton, C.T. et al. “Genomic reconstruction of an uncultured hydrothermal vent gammaproteobacterial methanotroph (family Methylothermaceae) indicates multiple adaptations to oxygen limitation”. In: *Frontiers in Microbiology* 23.6: 1425 (2015). doi: 10.3389/fmicb.2015.01425. URL: <https://doi.org/10.3389/fmicb.2015.01425>

K. Y.C performed laboratory work to prepare samples for sequencing

[8] Flytzanis, N.C. et al. “Archaeorhodopsin variants with enhanced voltage-sensitive fluorescence in mammalian and *Caenorhabditis elegans* neurons”. In: *Nature Communications* 15.5: 4894 (2014). doi:10.1038/ncomms5894.

K.Y.C. performed experiments.

## Table of contents

Acknowledgments.....	iii
Abstract .....	vi
Published Content and Contributions.....	viii
Table of Contents.....	ix
Chapter I: Introduction .....	1
1.1 Technology development for the study of biological systems.....	1
1.2 Tissue clearing for whole organs using the vasculature.....	2
1.3 Developing tissue clearing to render whole bone transparent .....	3
1.4 1.4 Characterizing adeno-associated viral capsids and engineering viral vectors .....	5
Chapter II: Whole-body tissue stabilization and selective extractions via tissue- hydrogel hybrids for high-resolution intact circuit mapping and phenotyping .	8
2.1 Summary.....	8
2.2 Building a PARS chamber .....	8
2.3 Standardized Procedure for PARS.....	10
2.4 Main Figures.....	12
Chapter III: Bone CLARITY: clearing, imaging, and computational analysis of osteoprogenitors within intact bone marrow .....	15
3.1 Summary.....	15
3.2 Introduction .....	16
3.3 Results.....	18
3.4 Discussion.....	24
3.5 Main figures.....	27
3.6 Supplementary Figures.....	34
3.7 Supplementary movie captions .....	45
3.8 Materials and methods .....	46
3.9 Supplementary methods .....	49
3.10 Additional information.....	54
Chapter IV: Engineered adeno-associated viruses for efficient and noninvasive gene delivery throughout the central and peripheral nervous systems.....	55
4.1 Summary.....	55
4.2 Introduction .....	56
4.3 Results.....	58
4.4 Discussion.....	64
4.5 Main Figures.....	68
4.6 Supplementary figures.....	78
4.7 Supplementary tables .....	91
4.8 Supplementary movie caption.....	93
4.9 Methods .....	94
4.10 Additional information.....	100
Bibliography .....	102

## *Chapter 1*

### INTRODUCTION

#### **1.1 Technology development for the study of biological systems**

Scientific discoveries are not only dependent on new and novel ideas or conceptual leaps that result in paradigm shifts, but also in advances to the technologies that make these steps possible. Tools that enable viewing or manipulation of biological systems have accelerated efforts in our understanding of complex biological processes by allowing researchers to visualize or perturb these events. Some of the most important and widely adopted biological tools are simple, elegant, and have revolutionized the way we perform research. One classic example of this is the discovery and usage of green fluorescent protein in *E. coli* and *C. elegans*. Since its publication, it has allowed researchers to peer into the inner workings of biological environments. Another example of this are the tools in optogenetics, where the use of light-activated bacterial proteins have allowed researchers to perturb biological circuits and dissect out their function with high spatial and temporal resolution in genetically defined populations of cells<sup>1</sup>. These examples highlight the importance of tool development in the study of biological systems. As the complexity of biology unravels, the demand for more advanced tools and techniques rises, and consequently, tool developers will need to adapt to this demand in order to further push discoveries with novel approaches. Limited to our current repertoire of tools, are those that allow for visualization of native three dimensional environments. Additionally, it is challenging to deliver transgenes into difficult-to-access areas, such as the cardiac ganglia, or broadly distributed networks, such as the myenteric nervous system, which limits our ability to extensively study these areas. Therefore, technological tools that overcome these limitations are in need.

The body of work summarized in this dissertation has been focused on developing tools in addressing a central question in biology: how different cell types interact with each

other and their native environment to form complex functional systems and networks. These tools include (i) developing tissue clearing methods to render tissue transparent for three-dimensional optical interrogations, (ii) expanding the adeno-associated viruses (AAV) toolbox by characterizing capsids for non-invasive widespread gene delivery throughout the CNS and PNS, (iii) generating or validating gene regulatory elements for cell type-specific expression of transgenes in non-transgenic animals, and (iv) developing viral vectors for sparse multi-color gene expression for single-cell morphology studies. Collectively, the developed tissue clearing methods, characterized novel viruses, and engineered viral vectors allow for single-cell resolution phenotyping in whole organs and will help further our understanding of complex biological systems and networks.

## **1.2 Tissue clearing for whole organs using the vasculature**

The ability to visualize biological tissue has long been appreciated by scientists. That is why tools like microscopes and microtomes have been indispensable in research. They allow us to view physical evidence of cellular components and processes in two-dimensional sections. But recently, with the growing ability to handle and process bigger data sets, more advanced microscopes being built (i.e., objectives with longer working distances and light-sheet microscopy allowing for faster acquisition times), and multidisciplinary approaches to tackle on scientific questions on a larger scale, it is no surprise that there is a trend to investigate biological tissue in three dimensions to better understand how cells interact with each other and their native environment. The brain is a prime example that represents the complexity of an organ and the need to study it in three dimensions. Neurons can project long distances and branch in many different directions to form intricate neural networks. However, the idea of rendering tissue transparent for three-dimensional optical access has been around for over a century. For example, solvent-based clearing methods, such as benzyl alcohol, was first described by Spalteholz over one hundred years ago<sup>2</sup>. Since then, many solvent-based variations have been published to render tissue transparent. These methods were able to produce samples that were novel at the time and allowed unprecedented optical access to study anatomy. However, major limitations to solvent-based clearing were the

inability to perform immunostaining, preserve native fluorescence, and risks damage to the tissue from the use of harsh chemicals<sup>3</sup>. With the ever growing number of transgenic reporter mouse lines and tools that genetically targeted cells with fluorescent proteins it was only a matter of time until researchers would come up with ways to render tissue transparent while preserving the ability to perform immunostaining and native fluorescence. The hurdle was crossed in 2013 when Chung K. et al. published the CLARITY technique<sup>4</sup> allowing the ability to immunostain and preserve native fluorescence. CLARITY utilized a hydrogel mesh that locked the proteins in their original positions while removing lipids from the tissue. Lipids are the main organic material that prevents imaging deep into tissue because of their light-scattering properties. However, a caveat to CLARITY was the use of an electric current to speed up the lipid removal process and penetration of antibodies. The use of this electric current caused the tissue to heat up and turn brown mainly due to the Maillard reaction. CLARITY also required expensive equipment, such as platinum wires and a commercial refractive index matching solution, which limited its wide adoption. Ultimately, researchers aimed to expand upon the CLARITY approach by avoiding the use of the electric current and also making the procedure more adoptable. In Chapter 2, I present tissue clearing work where I aimed to develop and optimize ways to perform rapid whole organ clearing of soft tissue using the vasculature, an idea published in the lab<sup>5</sup> but its procedure was not yet optimized. I standardized this procedure by detailing a step-by-step protocol and also outlined how to easily build an affordable clearing chamber with common and readily available parts found in a molecular laboratory. By using the vasculature, the use of an electric current could be avoided while still maintaining a rapid clearing time. In addition, small molecular dyes and antibodies could be perfused through to allow for immunostaining. This work is summarized in Chapter 2.

### **1.3 Developing tissue clearing to render whole bone transparent for single-cell phenotyping**

Since the publication of CLARITY and the proceeding clearing methods (uDISCO<sup>6</sup>, CUBIC<sup>7</sup>, PACT<sup>5</sup>, PARS<sup>5</sup>, ScaleSQ<sup>8</sup>), there was growing interest in adopting clearing

technology for the study of bone. There were two main reasons why this was important in relation to neuroscience. The first came from a paper in 2015 that detailed the discover of lymphatic vessels beneath the skull lining the dural sinuses<sup>9</sup>. Access to these areas beneath the skull through traditional means is difficult. They are readily damaged with even low amounts of force. Therefore, the study of such lymphatic vessels were difficult and for a long time overlooked. Methods that can study these areas of interest without disruption to the native environment could prove useful in fields like neuroimmunology and may help shed insight into how immune system dysfunction is associated with neurodegenerative diseases. The second reason came from the fact that bone contributes to whole-organism physiology. They regulate calcium levels, energy metabolism, and provide structural support<sup>10</sup>. Interestingly, there also exists a bone-brain axis that regulates appetite<sup>11</sup>. Therefore, the importance of bone to whole-organism physiology and accessing the delicate interfaces housing unique environments was a large motivation in developing tissue clearing technology to study bone. However, the challenge with bone is that it is a complex tissue; it contains both soft tissue (marrow) and hard tissue (minerals). At the time, tissue-clearing protocols were designed for soft tissue only, such as the brain. Consequently, a clearing protocol would need to be designed specifically for clearing bones. The goal was to develop a way to clear bone tissue while maintaining the native fluorescence of reporter mouse lines. This involved removing hard minerals from bones by decalcifying them with a chelating agent before lipid removal. In addition, since bone harbors a large amount of heme that autofluorescence, an amino alcohol was used to remove the heme. After I had developed a bone-clearing protocol, I collaborated with colleagues to use tissue clearing for bones to answer a biologically relevant question. I worked alongside Dr. Greenbaum in imaging intact bones with a custom-built light-sheet fluorescence microscope that minimized acquisition times compared to conventional point scanners microscopes. Together, we were able to image native fluorescence in whole bones at single-cell resolution to study a rare and non-uniform population of osteoprogenitor cells. I also worked with two young collaborators to build computation tools for semi-automated cell quantification of these rare and non-uniformly distributed osteoprogenitor cells. We found after administration of a novel

anabolic drug, we were able quantitatively detect a minute increase in the osteoprogenitor population of cells. We demonstrated that this minor increase may have been undetected using traditional methods until later time points due to sub-sampling errors from statistical estimation. This highlighted the sensitivity of our methods. The summary of this work is found in Chapter 3.

#### **1.4 Characterizing adeno-associated viral capsids and engineering viral vectors**

The process of introducing foreign DNA into a cell is one of the most fundamental techniques in biology. It allows researchers to manipulate a cell's phenotype for study or treatment. There are many ways to achieve gene transfer *in vivo*: electrically (electroporation), mechanically (microinjection), synthetically (nanoparticles), or biologically (viruses), to name a few. There is no gold-standard way, and each of the listed methods has its pros and cons. Nevertheless, the methods listed are well-established techniques and widely adopted in research. However, if there were a gold standard for *in vivo* gene transfer, it would probably be something that was: non-invasive, had a high transduction efficiency, with either broad or specific tropisms, have low immunogenicity, with the ability to deliver large-sized genomes, be cost efficient to produce, and allowed for long-term expression without host genome integration. With no surprise, various viruses have evolved from nature that are able to satisfy a moderate amount of these requirements, but not all. For example, herpes simplex virus HSV allows for large packaging of genomes and allows for fast expression in neuronal cells; however, HSVs can cause an acute inflammatory response<sup>12</sup>. Lenti viruses, on the other hand, have a modest packaging capacity and can transduce dividing and non-dividing cells, but generally contain impurities in their envelope, making them difficult to purify which causes an acute inflammatory responses<sup>13</sup>. Recombinant adeno-associated viruses (rAAVs) have a small packaging capacity compared to other viruses; however, their ability to provide safe and moderate levels of long-term expression in a wide array of cell types has made them an attractive gene delivery tool in basic research and clinical trials<sup>14</sup>. For example, in 2012, the first AAV gene therapy was approved by the European Commission for the treatment of lipoprotein lipase deficiency<sup>15</sup>.

This was a major milestone for viral gene therapy. Although AAVs are widely adopted in research and preclinical trials nowadays, their full potential is only beginning to unravel. Within the AAV platform, there is a lack of AAVs that can provide broad and non-invasive gene transfer to difficult-to-access environments (i.e., dorsal root ganglion, cardiac ganglia) or widely distributed organs (brain, enteric nervous system). In addition, the main delivery method of AAVs is through site-directed injections, which for most applications is suitable and desirable to restrict expression of a transgene. However, for applications, such as gene therapy in neurological diseases within the central nervous system (CNS), where likely multiple brain regions are affected, the current repertoire of AAVs would require numerous site-directed injection due to their poor transduction efficiency during systemic delivery. Furthermore, for basic research, the ability for widespread gene delivery could allow researchers to more rapidly test new molecular tools, such as voltage indicators, and refine those tools before dedicating the time and cost in generating a transgenic mouse line. In addition, AAVs that could deliver genetic tools into difficult-to-target areas could help in our studies in understanding these areas. From the already desirable properties of AAVs, it wasn't surprising that researchers focused on improving the versatility of the AAV platform. For example, researchers have engineered AAVs with more specific tropism to endothelial cells<sup>16</sup> in the CNS or to have enhanced transduction efficiency into the CNS<sup>17</sup> through systemic delivery. In 2016, Deverman et al. published AAV-PHP.B using a capsid selection method, called Cre recombination-based AAV targeted evolution (CREATE)<sup>18</sup>. AAV-PHP.B allowed for the ability to deliver genes across the central nervous system (CNS) non-invasively with a 40-fold higher transduction efficiency than the current standard, AAV9, at the time. Such AAVs may prove to be important for the treatment of neurological diseases where the replacement of a defective gene in the entire central nervous system would require widespread non-invasive gene delivery. However, much work is still required to improve these new novel AAVs. For example, AAVs with efficient non-invasive delivery into the peripheral nervous system (PNS) are lacking. These AAVs could be useful for the delivery of genes into areas that are difficult-to-target surgically, such as the dorsal root ganglia, for the study of pain. In addition, high viral payloads for efficient widespread transduction

through systemic delivery are still required. The ability to maintain a high transduction efficiency with a lower viral payload could be useful in several ways. The first is lowering viral payloads could help bypass neutralizing antibodies in the organism when high viral payloads are used. The second is more practical: lower viral payloads with high transduction efficiency will help in reducing the cost and time burden in viral production when scaling up for larger model organisms or preclinical trials. The viral work presented in this thesis is focused on the characterization of novel AAVs for efficient and noninvasive gene delivery throughout the CNS and PNS. The work is based off the CREATE method and uses it to further evolve the AAV-PHP.B capsid towards more efficient transduction in neurons at lower viral payloads. It also involves the characterization of a new capsid that showed high transduction efficiency towards peripheral sensory neurons. Viruses that can transduce sensory neurons widely distributed, such as those found in the enteric nervous system, are lacking, so this variant may prove useful in accelerating research in newly developing multi-system fields, such as the gut-brain axis<sup>19</sup>. In addition, I present my work in developing viral vectors for single-cell phenotyping in whole-organ systems. This was done by using our novel AAVs with engineered viral vectors I constructed to allow for sparse multicolor labeling of cells in a ubiquitous or cell type-specific manner. The idea of using multicolor for cell tracing and morphology studies is not novel. It was first published in 2007 by Livet J et. al, and afterwards, several iterations<sup>20</sup> were made to improve its approach. However, a major limitation to the approach was the lack of ability to modulate its density in labeling, something difficult to achieve with transgenic animals or site-directed injections of viruses. To overcome this, I worked on building and testing two different two-component systems using Flp recombinase or the tetracycline inducible system for modulating the density of labeling with our systemically deliverable viruses. I utilized this system to study cell morphology in the CNS and PNS. Finally, in order to further maximize the utility of AAVs, I also constructed or validated several different gene regulatory elements for cell type specific expression in non-transgenic animals. Collectively, this expansion of the AAV toolbox enables efficient and versatile gene manipulation in the CNS and PNS of transgenic and non-transgenic animals. This work is summarized in Chapter 4 of my thesis.

## Chapter 2

### Whole-body tissue stabilization and selective extractions via tissue-hydrogel hybrids for high-resolution intact circuit mapping and phenotyping

[6] Treweek, J.B. et al. “Whole-body tissue stabilization and selective extractions via tissue-hydrogel hybrids for high-resolution intact circuit mapping and phenotyping”. In: *Nature Protocols* 10.11 (2015), pp. 1860-1896. doi:10.1038/nprot.2015.122.

#### 2.1 Summary

To facilitate fine-scale phenotyping of whole specimens, we describe here a tissue fixation-embedding, detergent-clearing and staining protocol that can be used to transform whole organisms into optically transparent samples within 1–2 weeks without compromising their cellular architecture or endogenous fluorescence. The method perfusion-assisted agent release in situ (PARS) is optimized and standardized in this chapter for whole-body tissue clearing of soft tissue. It uses a hydrogel to stabilize tissue biomolecules during selective lipid extraction, resulting in enhanced clearing efficiency and sample integrity. Furthermore, the macromolecule permeability of PARS-processed tissue hybrids supports the diffusion of immunolabels throughout intact tissue, whereas RIMS (refractive index matching solution) grants high-resolution imaging at depth by further reducing light scattering in cleared and uncleared samples alike. The protocol and solutions outlined enable phenotyping of subcellular components and tracing cellular connectivity in intact biological networks.

#### 2.2 Building a PARS chamber

To perfuse PARS reagents through vasculature in a contained environment, we construct a PARS chamber using components that are readily found in most of the biological research laboratories (**Figure 2.1**). The necessary components of a PARS setup are as follows: a feeding needle catheter to deliver PARS reagents to vasculature; a perfusate catch-basin (pipette box) in which recirculating PARS reagents may pool once they exit the vasculature; Tygon tubing threaded through a peristaltic pump so that pooling

reagents may be collected from the catch-basin and recirculated back into a subject's vasculature; Luer-to-tubing couplers; and a Ziploc bag to contain the entire PARS chamber setup.

To construct the PARS chamber, drill two 1/8-inch holes into the front and one 1/8-inch hole into the left side wall of an empty 1,000- $\mu$ l pipette tip box. The holes are drilled just below the tip wafer (in **Figure 2.1**, the holes are ~2 cm below the top rim). Next, snap 1/8  $\times$  1/8-inch barbed connectors into each of the drilled holes. The outflow line will circulate solvents from the pipette box chamber to the 3-way stopcock. To join the outflow line to a 3-way stopcock, use a 10-cm piece of Tygon tubing and connect one end to the inner left side 1/8  $\times$  1/8-inch barbed connector and tape the other end to the inside bottom of the pipette tip box. Continue this line through a peristaltic pump by using a new piece of Tygon tubing, and connect one end to the outer left side 1/8  $\times$  1/8-inch barbed connector and then thread the tubing through a peristaltic pump. Next, join the free end to a three-way stopcock with a 3/32-inch barbed male Luer with locking nut. The inflow line circulates solvents from the three-way stopcock to the vasculature. To link the inflow line to the pipette tip box, use a piece of 15-cm Tygon tubing and connect one end to the outer right front 1/8  $\times$  1/8-inch barbed connector and the other end to the three-way stopcock with a 3/32-inch barbed male Luer with locking nut. To finish the inflow line, connect a piece of 75-cm Tygon tubing to the inner right front 1/8  $\times$  1/8-inch barbed connector. Next, coil the inflow line to the bottom of the pipette tip box. This will equilibrate inflowing solvent to the desired temperature before it enters the subject's vasculature. Tape the coiled tubing to the pipette tip box. To quickly circulate bubbles formed during the changing of solutions without disconnecting the inflow line and for use of bubbling nitrogen gas into the solution during the polymerization step, a line linking the outflow line back to the pipette tip box is connected by joining a piece of 15-cm Tygon tubing to the three-way stopcock with a full-thread 3/32-inch barbed female Luer to the outer left front 1/8  $\times$  1/8-inch barbed connector. This line is continued inside the pipette tip box and taped to the bottom. To finish the chamber, thread the inflow line through the top-left corner of the tip wafer and connect it to a feeding tube with a 1/8-inch barbed male slip Luer. As a

forewarning, SDS and salt precipitate will begin to accumulate within these narrow lines over time. It is important to flush the lines (e.g., with ddH<sub>2</sub>O) between subjects, and to replace occluded lines with new Tygon tubing (e.g., after every few subjects).

During hydrogel polymerization, the chamber must be enclosed inside a Ziploc freezer bag. To do this, disconnect the outer Tygon tubing that connects to the barbed connectors of the pipette tip box, and puncture three holes into the Ziploc bag to accommodate the 1/8 × 1/8-inch barbed connectors. Reconnect the Tygon tubing to their original 1/8 × 1/8-inch barbed connector. To connect a vacuum line to this bagged PARS box for withdrawing oxygen, tape a female Luer tee onto the lid of the pipette box and puncture one hole through the Ziploc. Finally, make the Ziploc airtight by placing clay around the punctured regions in the Ziploc.

As a final note, a 1,000- $\mu$ l tip box has a volume of ~750 ml. Thus, during hydrogel polymerization and during clearing, 200–300 ml of solution may be placed in the pipette box for recirculation without risk of the pipette box overflowing or the solution splashing out during its transport. Similarly, to conserve reagents during PARS clearing and immunostaining of smaller samples, a 200- $\mu$ l tip box may be used to construct the PARS chamber; only 100 ml of reagent is necessary to fill such a chamber about one-third full (**Figure 2.1**).

### **2.3 Standardized Procedure for PARS**

After polymerization, wash the perfusion lines with 1× PBS, and then replace the wash buffer with 8% SDS-PBS (pH 7.5) clearing buffer. This procedure can be accomplished easily by removing the PARS chamber from the shaking water bath (optional), by turning off the pump, by removing excess initiator buffer from the PARS chamber and by replacing it with 100 ml of 1× PBS to perform the wash. Circulate the wash buffer through the sample for 10 min. Afterward, replace the buffer with 8% SDS-PBS (pH 7.5). Place the chamber back into the Ziploc bag and into the 37–42 °C water bath. Allow the SDS clearing buffer to recirculate through the system for 24 h. Perform a buffer exchange with fresh 8% SDS-PBS

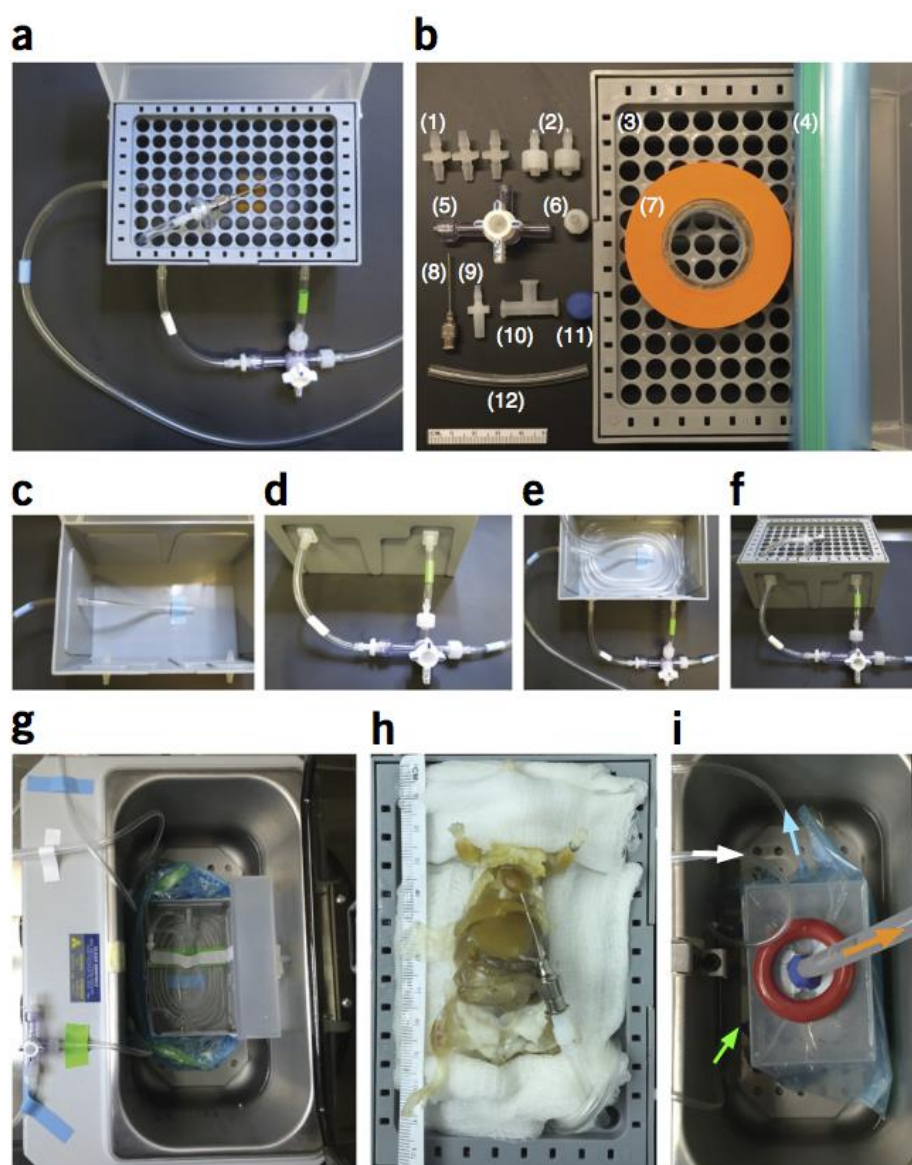
(pH 7.5), by clearing the buffer daily until the recirculated fluid is no longer yellowish, after which the SDS solution can be refreshed less frequently (every 48–72 h).

Check on the clearing progress daily. Add additional SDS buffer to the PARS chamber if necessary, as depending on how well the Ziploc bag is sealed around the perfusion tubing, some buffer may evaporate over time.

The sample can be continuously perfused for up to 2 weeks until all desired organs have cleared, even if some organs appear clear within the first 24–48 h. Alternatively, if all but one or two organs appear sufficiently transparent after a few days, one may proceed directly to Step 7C(iv) to flush SDS from tissue, and then excise all organs. The one or two semiopaque excised organs are transferred into 8% (wt/vol) SDS to finish clearing via PACT, whereas the organs that cleared more rapidly are immediately promoted to passive immunostaining (optional) and mounting without further delay.

For samples that will be immunolabeled after clearing, the hydrogel matrix will be required to support the cleared tissue for several rounds of washing and multiday incubations with gentle shaking. If the already delicate tissue-hydrogel matrix seems precariously fragile after clearing (this usually only occurs with thin-sectioned tissue), it is advisable to repeat the hydrogel embedding and polymerization steps. This will stabilize tissue architecture during immunolabeling, prevent tissue loss or disintegration and counteract expansion in mounting medium. For cleared samples that will not undergo any immunohistochemical labeling steps before imaging. The ending results are in **(Figure 2.2)**

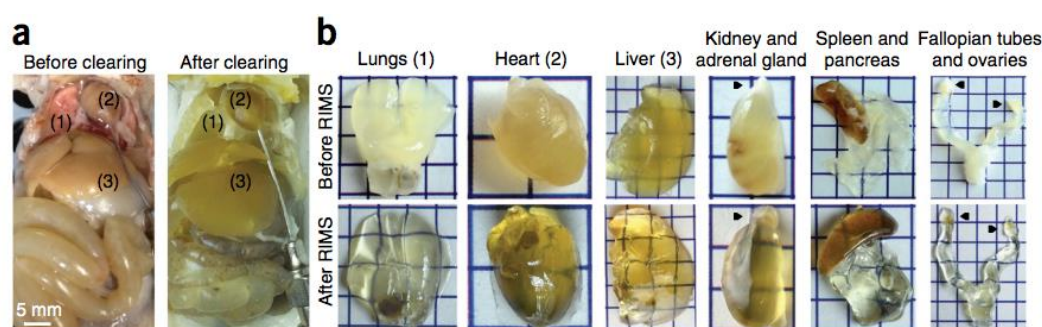
## 2.4 Main Figures



**Figure 2.1.** Assembling and working with the PARS chamber. (a) A completed PARS chamber used for whole-body tissue clearing. (b) Individual parts to build a PARS chamber: (1) three  $1/8 \times 1/8$ -inch barbed connectors, (2) two  $3/32$ -inch barbed male Luers with locking nut, (3) a 1,000  $\mu$ l pipette tip box, (4) a 1-gallon Ziploc freezer bag, (5) a three-way stopcock with Luer lock, (6) a  $3/32$ -inch barbed female Luer with full tread, (7)

a roll of lab tape, (8) a 22-G  $\times$  1-inch gavage needle, (9) a 1/8-inch barbed male slip Luer, (10) a female Luer tee with locks, (11) clay and (12) Tygon E-lab tubing. Ruler shown is 5 cm in length. (c) Three 1/8-inch holes are drilled into the pipette tip box: two into the box front and one into its side, all  $\sim$ 2 cm below the top rim of the box. The three 1/8  $\times$  1/8-inch barbed connectors are placed into the drilled holes. To connect the outflow line (blue tape bands on outflow line tubing), a piece of Tygon tubing is connected from the bottom inside of the pipette box to the single 1/8-inch barbed connector that was inserted through the box side. (d) To continue the outflow line, a second, longer piece of blue-taped tubing is attached to the outer fitting of this same barbed connector (on the outside of the pipette tip box side), and then the other end of this tubing is threaded through the peristaltic pump, pulled back over toward the pipette box and finally connected to a three-way stopcock with a 3/32-inch barbed male Luer with locking nut (rightmost blue-banded tubing in d). To form the inflow line, a short length of tubing (green tape band) is used to connect the three-way stopcock to the front right 1/8-inch barbed connector of the pipette box. The solute flushing line and nitrogen bubbling line, which are subserved by the same tubing (white tape band), are formed by another short length of tubing that joins the third port of the stopcock to the front left 1/8-inch barbed connector. (e) The inflow line is continued inside the pipette box, with the tubing coiled several times around the base of the box so that the solute will be reheated before it passes through the feeding gavage into the subject. The solute flushing line and nitrogen bubbling line is continued inside the pipette tip box and taped to the bottom of the chamber (not shown). (f) The tip of the coiled inflow line tubing is threaded up through the tip wafer (see bird's-eye view of threaded wafer in a) and connected to a 22-G  $\times$  1-inch gavage needle with a 1/8-inch barbed male slip Luer. The gavage needle is secured with a short loop of Tygon tubing ( $\sim$ 90 mm) threaded through two holes of the wafer. (g) During the polymerization step, the chamber is placed into a 37  $^{\circ}$ C water bath and sealed in a Ziploc bag. The tubing is attached to the chamber with three 1/8  $\times$  1/8-inch barbed connectors punctured through the Ziploc bag. The Tygon tubing is reconnected from the outside of the bag and surrounded with clay to make an airtight seal. (h) The animal is placed onto the pipette tip box, and the 22-G  $\times$  1-inch gavage needle is

used to catheterize the heart. (i) The chamber is placed into a 37 °C water bath. A female Luer tee, which is taped onto the lid of the pipette tip box, is punctured through the Ziploc bag, and this joint is sealed with clay to ensure an airtight seal. Finally, to accelerate polymerization, a vacuum line is connected to the female Luer tee to remove oxygen (orange arrow), and a nitrogen gas line (white arrow) is connected to the 1/8-inch barbed connector to deliver a steady flow of nitrogen into the bagged system. The solute is continually circulated through the animal from the outflow line (blue arrow, which also indicates the direction of flow through blue-taped tubing) and inflow line (green arrow, which also indicates the direction of flow through green-taped tubing).



**Figure 2.2.** Whole-body clearing of mice with PARS. (a) A4P0-hybridized organs shown before the start of clearing (left) and after 5 d of clearing with 8% SDS-PBS (pH 8.5) and overnight washing with 1× PBS at pH 7.5 (right). Numbers correspond to the extracted organs in b. (b) Extracted organs from the cleared mouse in a, pictured before (top) and after (bottom) RIMS incubation for 3 d. Black pointers correspond to the adrenal gland on the kidney and to the ovaries on the fallopian tubes. Each square represents 0.5 cm<sup>2</sup>. Rodent husbandry and euthanasia conformed to all relevant governmental and institutional regulations; animal protocols were approved by the Institutional Animal Care and Use Committee (IACUC) and by the Office of Laboratory Animal Resources at the California Institute of Technology.

## Chapter 3

### **Bone CLARITY: clearing, imaging, and computational analysis of osteoprogenitors within intact bone marrow**

[2] Greembaum, A. et al. “Bone CLARITY: Clearing, imaging, and computational analysis of osteoprogenitors within intact bone marrow”. In: *Science Translational Medicine* 387.9 (2017). URL: <http://doi.org/10.1126/scitranslmed.aah6518>.

#### **3.1 Summary**

Bone tissue harbors unique and essential physiological processes such as hematopoiesis, bone growth, and bone remodeling. To enable visualization of such processes at the cellular level in an intact environment, we developed “Bone CLARITY”, a bone tissue clearing method. We used Bone CLARITY and a custom-built light-sheet fluorescence microscope (LSFM) to detect the endogenous fluorescence of Sox9-tdTomato<sup>+</sup> osteoprogenitor cells in the tibia, femur, and vertebral column of adult transgenic mice. To obtain a complete distribution map of these osteoprogenitor cells, we developed a computational pipeline that semi-automatically detects individual Sox9-tdTomato<sup>+</sup> cells in their native 3D environment. Our computational method counted all bone cells without relying on sampling techniques and displayed increased precision when compared with traditional stereology techniques for estimating the total number of rare cells. We demonstrate the value of the clearing–imaging pipeline by quantifying changes in the population of Sox9-tdTomato labeled osteoprogenitor cells following sclerostin antibody treatment. Bone tissue clearing is able to provide fast and comprehensive visualization of biological processes in intact bone tissue.

### 3.2 Introduction

The mammalian skeletal system consists of numerous bones of varying shapes and sizes that provide support to the body and protect internal organs from external physical stress<sup>21,22</sup>. Different bone types harbor specialized physiological processes that are key for proper development and survival of the organism, such as replenishment of hematopoietic cells, growth, and remodeling of the bone during healthy and diseased states<sup>23,24 25,26</sup>. Traditionally, these processes have been investigated through methods that provide zero- or two-dimensional information, such as fluorescence activated cell sorting (FACS) or analysis of histological sections. Quantitative 3D data of geometric features such as volume and number of cells, can be obtained from histological sections with unbiased stereological methods. Although statistically robust, such methods are labor-intensive and provide no visualization of the 3D structures. The need for methods that provide 3D information to study the bone has long been recognized. While methods such as serial sectioning and milling are valuable tools for understanding the structure of bone at the tissue level, these are destructive techniques that do not provide information at the cellular level and cannot be easily combined with other methods such as immunohistochemistry to characterize cellular processes<sup>27,28</sup>.

CLARITY (Clear Lipid-exchanged Acrylamide-hybridized Rigid Imaging / Immunostaining / in situ-hybridization-compatible Tissue hYdrogel) was originally developed for soft tissues such as the brain<sup>4</sup> recently there has been a surge in optical clearing methods for a variety of applications, for example profiling of tumor biopsies and brain tissue<sup>4,5,29-33</sup>. Bone is a more complex histological sample owing to its hard (mineral) and soft (bone marrow) tissue, and accordingly osseous tissue has remained a challenge, despite some promising developments reviewed in<sup>32,34</sup>. One method for investigating intact bones and their 3D microenvironments at sub-micron resolution is to render the tissue optically transparent. Bone transparency can be accomplished by coupling refractive index (RI) matching reagents with removal of minerals and lipids that scatter light. Most bone clearing literature is based on solvent clearing methods<sup>35-38</sup>. These methods focus on RI matching and solvation of some lipids, but do not remove minerals<sup>32,34</sup>. In general, these

solvent-based clearing methods have achieved an imaging depth of approximately 200  $\mu\text{m}$  using two-photon microscopy<sup>36</sup>. Murray's clearing method was recently modified to clear bisected long bones and achieved an imaging depth of approximately 600  $\mu\text{m}$  with confocal microscopy<sup>39</sup>. Despite these advances, manipulation and subsampling of the bone is required for deep imaging, thus disrupting the intact bone architecture. A key limitation of Murray's clearing method and its variants is that they quench endogenous fluorescence, minimizing their application with transgenic fluorescent reporter lines, which are used to highlight key cell populations within the bone and marrow. Consequently, there is a need for a clearing method that maintains the intact bone structure, preserves endogenous fluorescence, and allows deeper imaging within intact bone without manipulation for subsampling.

In our prior work<sup>33</sup> we noticed that decalcification (deCAL) of bones by EDTA could in principle expand CLARITY applications to osseous tissues but we only achieved modest optical access (200–300  $\mu\text{m}$ ). In this work we introduce deCAL CLARITY for bones, a specialized protocol that incorporates continuous convective flow during the clearing process, amino alcohol to minimize tissue autofluorescence<sup>40,41</sup>, and an imaging procedure that minimizes refractive index variations in light-sheet microscopy. These improvements allowed us to achieve whole-bone clearing with an imaging depth of up to approximately 1.5 mm, while maintaining native tdTomato fluorescence and a signal-to-noise ratio (SNR) that permitted detection and 3D placement of single cells. We present a comprehensive platform based on whole bone clearing, light sheet imaging with a custom-built microscope, and dedicated computational methods for counting fluorescently labeled cells. We use this trio of methods to visualize and quantify the total number of osteoprogenitors contained within a volume of mouse bone and map their 3D spatial distribution in response to a sclerostin antibody (Scl-Ab), a bone-forming agent.

### 3.3 Results

#### *Bone CLARITY renders intact bones transparent while preserving endogenous fluorescence*

We developed and applied a bone clearing method to render the tibia, femur, and vertebral column of mice transparent for light-microscopy investigation (Fig. 1A). The key steps of the sample preparation including tissue clearing and autofluorescence removal are outlined in Fig. 1B and fig. S1. The bone is decalcified to increase light and molecular penetration through the tissue while leaving a framework of bone matrix with similar structural characteristics to dense fibrous connective tissue. DeCAL CLARITY employs an acrylamide hydrogel to support the tissue structure and minimize protein loss prior to the delipidation step. The detergent sodium dodecyl sulfate (SDS) is used to remove lipids in order to minimize their light scattering effects. We observed high autofluorescence in the bone marrow, one of the primary sites of heme synthesis. As heme is strongly autofluorescent, in the final step of the process we used the amino alcohol N,N,N',N'-tetrakis(2-hydroxypropyl)ethylenediamine<sup>40,41</sup> to remove heme, which minimized marrow autofluorescence by approximately 3-fold (fig. S2). All of the above clearing stages were conducted on a temperature controlled stir plate that provides continuous convective flow (fig. S3, auxiliary design file). This accelerates and improves the clearing process for entire organs compared with passive clearing<sup>33</sup>. The samples did not change size during the clearing process (fig S4).

The imaging of intact bone tissue using a point scanning method, for example using a confocal microscope, poses an operational challenge because acquisition times are prolonged, with concomitant photobleaching<sup>42,43 44</sup>. To address this, we built a custom light-sheet fluorescence microscope (LSFM) with confocal slit-detection (fig. S5)<sup>33 45</sup>. To modify the LSFM to image bones, an additional light-sheet path was added to illuminate the bones from two opposite directions. This additional light path improved optical access to areas containing a high percentage of cancellous bone, which typically scatter light more extensively and consequently lower the SNR. Given a particular depth scan, the sample is

illuminated by only one of the two light-sheet paths, whichever provides the better contrast. We use only one illumination path per depth scan because there is always one illumination direction that scatters less light and thus provides superior image quality. To minimize RI mismatch between the objective lens and the bones, the samples are directly immersed in the immersion chamber without the use of a quartz cuvette to hold the sample. If RI variations still persist, the position of the detection objective is changed for each tile and specific depth along the Z-scan to mitigate any resulting out-of-focus aberrations. The LSFM captures images at a frame rate of 22 frames per second (16-bit depth) and produces 0.176 GB of imaging data per second. Large datasets are thus acquired for each imaged bone (50–500 GB). In order to manage these large datasets, we designed a computational pipeline that includes image stitching, automatic detection of individual cells, and volume of interest (VOI) rendering for analysis (Fig. 1C).

To validate the protocol, we applied our clearing and imaging method to locate progenitor cells in the long bones and vertebrae of transgenic reporter mice. A Sox9CreER transgenic mouse line was used in which, upon tamoxifen injection, multipotent osteoblast and chondrocyte progenitor cells express tdTomato<sup>21,46</sup>. We visualized the endogenous fluorescence of Sox9<sup>+</sup> cells using Bone CLARITY (Fig. 1D and E). Quantification of the imaging depth in different regions of the tibia, femur, and vertebral body showed that we were able to image through the diaphysis of the femur (movie S1) and tibia (movie S2) and the entire vertebral body (movie S3). Furthermore, we were able to reliably detect Sox9<sup>+</sup> cells up to approximately 1.5 mm deep into the bones (fig. S6). Collectively, deCAL CLARITY coupled with LSFM and the data-processing pipeline is an effective clearing, imaging, and data processing protocol to investigate intact mouse bones.

### ***Semi-automated computational pipeline quantifies Sox9<sup>+</sup> cells in mouse tibia and femur***

To demonstrate that biological environments can be observed and quantified with Bone CLARITY, we first counted Sox9<sup>+</sup> cells in the tibia and femur. We found that Bone

CLARITY allows for detection and quantification of individual Sox9<sup>+</sup> cells in 3D (Fig. 2A), some of which appear associated with small vessels. Owing to the large VOI (Fig 2A, gray surface), we created a semi-automated cell-detection algorithm (Fig. 2B and fig. S7). The algorithm divides the 2D images that make up the Z-stack (or depth scan) into  $N$  small overlapping regions, and then adaptively thresholds the 2D images into binary images on the basis of the local mean and standard deviation of fluorescence. The resulting binary 2D images are further subject to morphological operations to eliminate noise and discontinuity within a cell. All of the 2D binary images are then combined into a 3D matrix. From the 3D matrix, only volumes that fit the properties of a Sox9<sup>+</sup> cell are maintained, thus avoiding erroneous counting of large blood vessels or small autofluorescence artifacts. All cells outside the user-defined VOI (Fig. 2A) are discarded. The cell candidate centroid locations are then imported to 3D visualization software for quality control. During the quality control stage, the annotator reviews the entire 3D volume and corrects the automatic results by marking false negative cells and omitting false positive cells. Therefore, at the end of the quality control stage, the cell counts are equivalent to cell counts that are performed manually. According to our experiments, a fully automatic pipeline only achieves on average 52% sensitivity and 36% precision. It should be noted that removing false positives is a faster operation in the 3D visualization software than adding false negatives, therefore the value of sensitivity outweighs the need for precision.

We quantified the number of fluorescently labeled Sox9 cells in the femur and tibia of Sox9CreER mice versus two control groups: Sox9CreER mice without tamoxifen administration and wild-type mice without the transgenes (Fig. 2C). The analyzed volume for each group was comparable (fig. S8). The cell counts in the Sox9CreER group without tamoxifen administration can be attributed to leakage of the reporter line because there was no visible expression of tdTomato in the wild-type control samples<sup>47</sup>. Qualitatively, this leakage can be observed in fig. S9, where the expression of the tdTomato signal is compared between the experimental group and the two control groups. We next quantified the cell distribution as a function of distance from the periosteal surface along the longitudinal plane (schematic in Fig. 2D). Based on our analysis, most of the cells in the

diaphysis reside adjacent to the endocortical surface, with mean distances from the periosteal surface of approximately 136.9  $\mu\text{m}$  and 143.6  $\mu\text{m}$  for the tibia and femur, respectively (Fig. 2E). This result supports similar findings in<sup>48</sup> and validates deCAL CLARITY as a reliable method to resolve and quantify individual cells in intact bone and marrow spaces. It should be noted that Bone CLARITY is not limited to visualize cell populations in transgenic animals only and that antibody staining is also feasible (fig. S10). However, for maximum penetration of the antibody into the bone, it is recommended to bisect the bone prior to the clearing process.

### ***Bone clearing can complement section-based stereology***

Design-based stereology is the gold standard method to quantify total cell numbers and densities in organs while preserving spatial information<sup>49</sup>. Stereology relies on statistical sampling methods. Systematic uniform random sampling (SURS) is a frequently used sampling method that efficiently reduces the variance of the estimate compared with random sampling. SURS obtains histological sections from an organ to reduce the amount of tissue for analysis. SURS samples at regular uniform intervals with the first sample collected at a position random within the first interval<sup>50,51</sup>. The number of cells in a tissue volume is a zero-dimensional geometric feature, and to avoid bias based on the cell size or shape, a probe based on two thin physical sections or optical planes separated by a known distance (disector), is used to accomplish quantification<sup>52,53</sup>. Cells are typically counted in a known fraction of the organ, which allows for an estimation of the total number of cells in the entire organ<sup>54</sup>. The spatial distribution of cells can also be obtained using second-order stereological methods<sup>55</sup>.

Applying stereological methods to cleared organs offers notable time-saving because clearing precludes the need for labor-intensive sectioning. Using the Bone CLARITY method and our 3D counting scheme, we investigated the variability in stereology estimates as a function of the number of slices for both SURS and simple random sampling in stereology experiment simulations. Figure 3A shows one

representative selection of  $N$  random uniformly spaced sections for a simulated SURS experiment in the femur, tibia, and vertebral body. Fig. 3B shows the estimates for the cell count of an entire VOI based on  $N$  sampled slices, and Fig. 3C illustrates the coefficient of variation of simulated stereology experiments that were conducted with  $N$  slices. To estimate the coefficient of variation for each  $N$  representative slices, 5 simulated stereology experiments were conducted. In these simulations, as expected, variance decreased rapidly with increasing number of slices for femur, tibia, and vertebrae. Precision of the stereological estimate would likely be improved with proportionator sampling, a form of non-uniform sampling better suited to rare structures<sup>56</sup>. Therefore, the 3D counting method offers several advantages for quantifying rare cellular populations: the ability to detect subtle changes that might be overlooked because of sampling variance, elimination of the need for sectioning, and 3D visualization.

### ***Sclerostin antibody increases the number of Sox9<sup>+</sup> cells in the vertebral column***

We next applied deCAL CLARITY to the vertebral column, a bone that is highly susceptible to fractures due to osteoporosis<sup>57</sup> and whose complex geometry is particularly difficult to probe with traditional sectioning-based methods. We cleared and imaged the vertebrae from mice, and focused on the 4<sup>th</sup> lumbar vertebral body (L4), which is dense, opaque, and predominately composed of cancellous bone. Fig. 4, A and B show schematics of lateral and transverse cuts from a representative vertebra. Fig. 4B also shows the approximate locations (dashed lines) of the digital sections (30  $\mu$ m thickness) that are shown in Fig. 4C. The lateral processes that extrude from the vertebral body can be seen in the horizontal edges of the 650  $\mu$ m section. These processes were excluded from the cell counts, as illustrated by the representative VOI (Fig. 4D). Additionally, careful attention was given to ensure the exclusion of the intervertebral discs and cartilage endplates from the VOI, as they are both primarily populated by chondrocytes. These chondrocytes express Sox9 and are tdTomato<sup>+</sup> after tamoxifen administration, and thus require careful exclusion for accurate cell counts of fluorescently marked osteoprogenitors. The

quantification of vertebral Sox9-tdTomato cells after tamoxifen administration versus the two control groups, Sox9CreER mice without tamoxifen administration and wild-type mice, can be seen in Fig. 4E. Again, we observed that the Sox9CreER transgenic animals without tamoxifen administration display mild leakage of tdTomato expression (Fig. 4E and fig. S9). Overall, similar to the cell distribution results presented for the tibia and femur, the Sox9<sup>+</sup> cells were primarily located adjacent to the endocortical surface (Fig. 4, F and G), with a mean distance of 138.8  $\mu\text{m}$  and 137.6  $\mu\text{m}$  for the tamoxifen-negative (Tam.<sup>-</sup>) and tamoxifen-positive (Tam.<sup>+</sup>) groups, respectively.

We next tested the effects of sclerostin antibody (Scl-Ab) on the total number of Sox9<sup>+</sup> cells in the vertebral column. Sclerostin is an extracellular inhibitor of the canonical Wnt signaling pathway and is highly expressed in osteocytes. Inhibition of sclerostin leads to activation of canonical Wnt signaling in the osteoblast lineage resulting in a rapid but transient marked increase in osteoblast number and bone formation<sup>58</sup>. Although the early dramatic increase in osteoblast number is considered to be at least in part due to activation of bone lining cells<sup>59</sup>, the contribution of osteoprogenitors to this early increase in osteoblast number is unclear. While a decrease in osteoprogenitors has been demonstrated to be associated with the attenuation of bone formation that occurs with long-term Scl-Ab treatment in rats, stereological methods have not detected effects on progenitor number coincident with the maximal increase in osteoblast number. In order to gain insight into the acute effects of Scl-Ab on osteoprogenitors, we conducted an experiment as outlined in Fig. 4H. On day one, Sox9CreER mice were injected with Scl-Ab (100 mg/kg subcutaneously). Four days afterwards, we provided a second dose of Scl-Ab along with tamoxifen to label Sox9<sup>+</sup> cells with tdTomato. We euthanized the animals five days later and cleared the bones using Bone CLARITY. Indeed, after performing blinded quantification of Sox9<sup>+</sup> cells, we observed an increase in the total number of Sox9<sup>+</sup> cells in the vertebral body nine days after initial Scl-Ab treatment versus the vehicle control group (Fig. 4I). We did not observe any significant changes to the vertebra body volume (Fig. 4J) between the experimental and control groups at this early time point. These results demonstrate that there is an increase in osteoprogenitors that are likely recruited to the bone

surface to contribute to the increase in osteoblast number. The distribution of cells as a function of distance from the surface can be seen in Fig. 4K, with mean distances of approximately 99.9  $\mu\text{m}$  and 137.8  $\mu\text{m}$  for the vehicle and treated groups, respectively.

### 3.4 Discussion

In the bone remodeling process, bone health is maintained through continuous cycles of bone resorption by osteoclasts and bone formation by osteoblasts. Imbalances in these physiological processes can lead to various bone diseases such as osteoporosis, which affect millions of people in the United States alone<sup>57,60</sup>. In order to gain better insight into potentially effective treatments for osteoporosis, it is imperative to study the physiological processes that occur in healthy and diseased bone and understand the molecular and cellular mechanisms within the 3D microenvironment. We demonstrate that the Bone CLARITY technique renders the tibias, femurs, and vertebral bodies of mice optically transparent while preserving bone morphology and an endogenous fluorescent reporter signal. In addition to matching the RI of the tissue, Bone CLARITY also removes minerals and lipids, thus enabling us to reconstruct a whole vertebral body as well as the entire diaphysis from the tibia and femur.

For maximum impact, clearing and imaging platforms need to be easy to use and scalable. The trio of methods presented here has a few limitations in its current rendering that could be improved in the future, such as the addition of antibody staining to the clearing workflow, reducing processing time and overcoming the technical barriers of acquiring and analyzing big datasets. Antibody staining of an intact bone is challenging owing to poor penetration by relatively large intact antibodies. In order to improve antibody penetration, we bisected and cleared half of a tibia and femur. In the bisected and cleared bones, the antibody penetrated up to 400  $\mu\text{m}$  from the bone surface and showed high specificity and SNR (fig. S10). The use of small-molecule staining methods (e.g. single-molecule hybridization chain reaction<sup>61</sup>) could allow the labeling of intact bones while achieving reliable staining depths. Meanwhile the ability to retain and detect endogenous fluorescence has proven highly enabling although the gentle clearing reagents used in

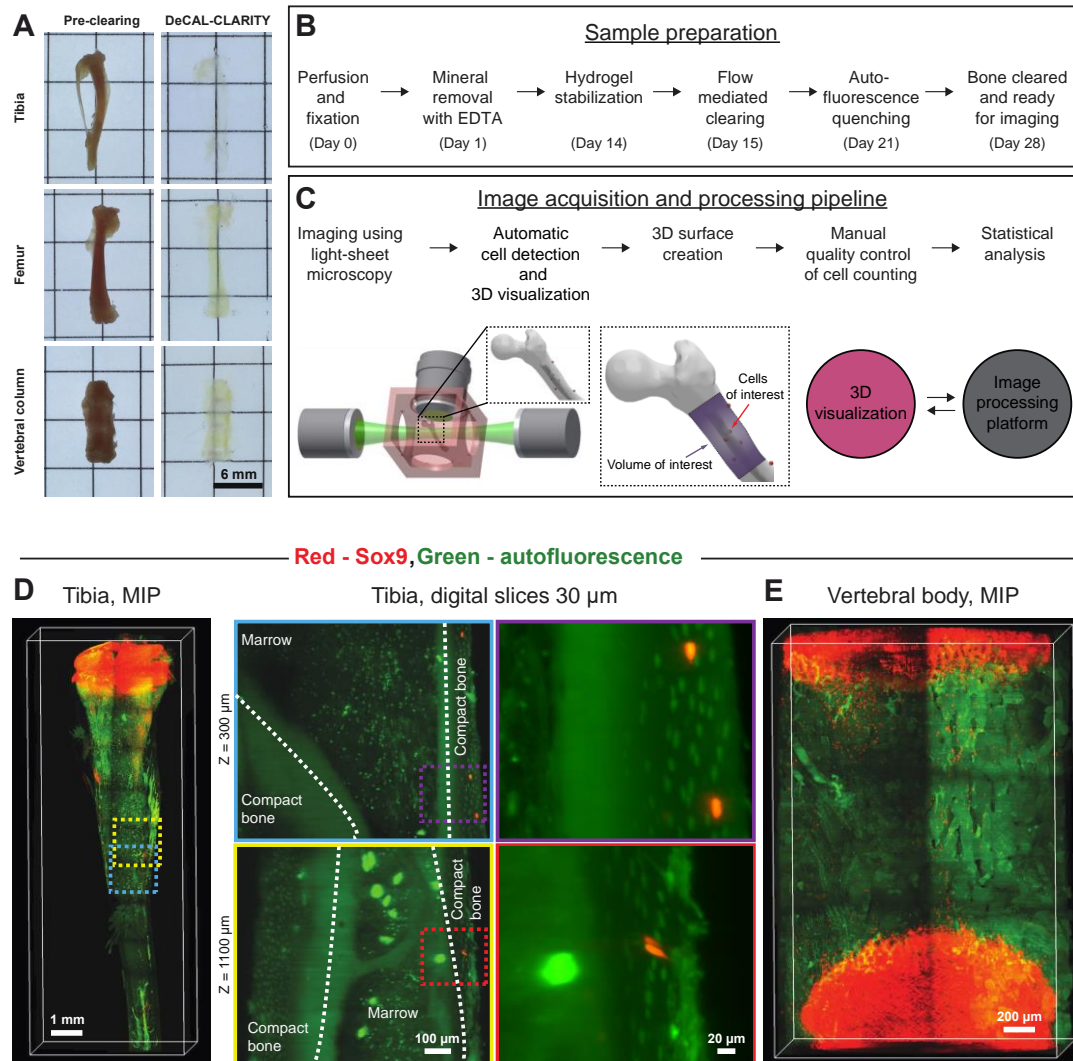
deCal CLARITY do introduce a compromise in processing time (28 days). Utilizing faster decalcification agents, such as formic acid, might shorten the current decalcification time of 14 days albeit the fluorescent proteins generated by the reporter genes used in this study might lose their fluorescence under acidic conditions<sup>42,62</sup>. Delipidation is also a lengthy stage in Bone CLARITY but necessary to reduce scattering from lipids present not only in the mineralized bone tissue, but highly abundant within the bone marrow. Consequently, delipidation ensures high-quality optical access deep in the bone. Although the use of customized microscopy and software might not be easy to implement in a non-technical setting, commercial LSM systems with a streamlined user interface and associated software are rapidly evolving to support the types of application described here. For data processing, we found that fully automated cell-detection algorithms were difficult to apply within cleared tissue. SNR variations arising from non-uniform illumination and fluorescence detection within the bone resulted in unsatisfactory precision, necessitating manual quality control. The simple automated tools that we developed are fast and adaptive and in general are able to save annotator time while improving precision and reducing error; however, more work is required to achieve a reliable, fully automated algorithm for cell counting. In general, data handling, visualization, and analysis would benefit from individual developers sharing their code in an open-source environment, which would allow the scientific and medical community to efficiently customize software relevant for the application at hand.

Bone CLARITY enables clearing of mouse bones while retaining the integrity of the bone marrow and endogenous fluorescence. Following clearing, we achieved 3D reconstructions of a vertebral body and long bone diaphyses, and attained an imaging depth of approximately 1.5 mm for the epiphysis of long bones. Using a computational pipeline to process large datasets and detect single cells in bone, we mapped the spatial distribution of osteoprogenitor cells. Additionally, we demonstrated the advantages of 3D methods in estimating rare cell populations that are not readily amenable to sampling by traditional stereology methods. It should be noted that combining traditional stereology methods with tissue clearing techniques can be advantageous, especially in cases where quantification of

complicated structural elements cannot be done automatically. Therefore, subsampling is advantageous for manual quantification within a reasonable timeframe. Finally, to further demonstrate the utility of our clearing method, we treated a cohort of adult reporter mice with a sclerostin antibody, a bone-forming agent, for 9 days. Previous stereological studies in rats treated with Scl-Ab for 8 days revealed a dramatic increase in total osteoblast number in the vertebrae coincident with increased bone formation, but no significant effect on bone progenitor numbers<sup>58</sup>. We found that, after 9 days of treatment in mice, the total number of osteoblast progenitor cells increased by 36% compared with the control group. This result was not surprising based on the literature<sup>58</sup>, but it has been challenging to demonstrate using stereological methods given the rarity of osteoprogenitor cells, particularly in the vertebrae. This underscores the greater sensitivity of our clearing, imaging, and data-processing protocol for quantifying rare cell populations as well as using lineage tracing to mark progenitors, as opposed to immunophenotyping in tissue sections. Overall, continued developments in tissue clearing<sup>6,32,34</sup>, imaging, and data analysis can facilitate translational research that will provide insight into the efficacy and safety of new bone-modulating drugs by profiling their effects on progenitor cell populations.

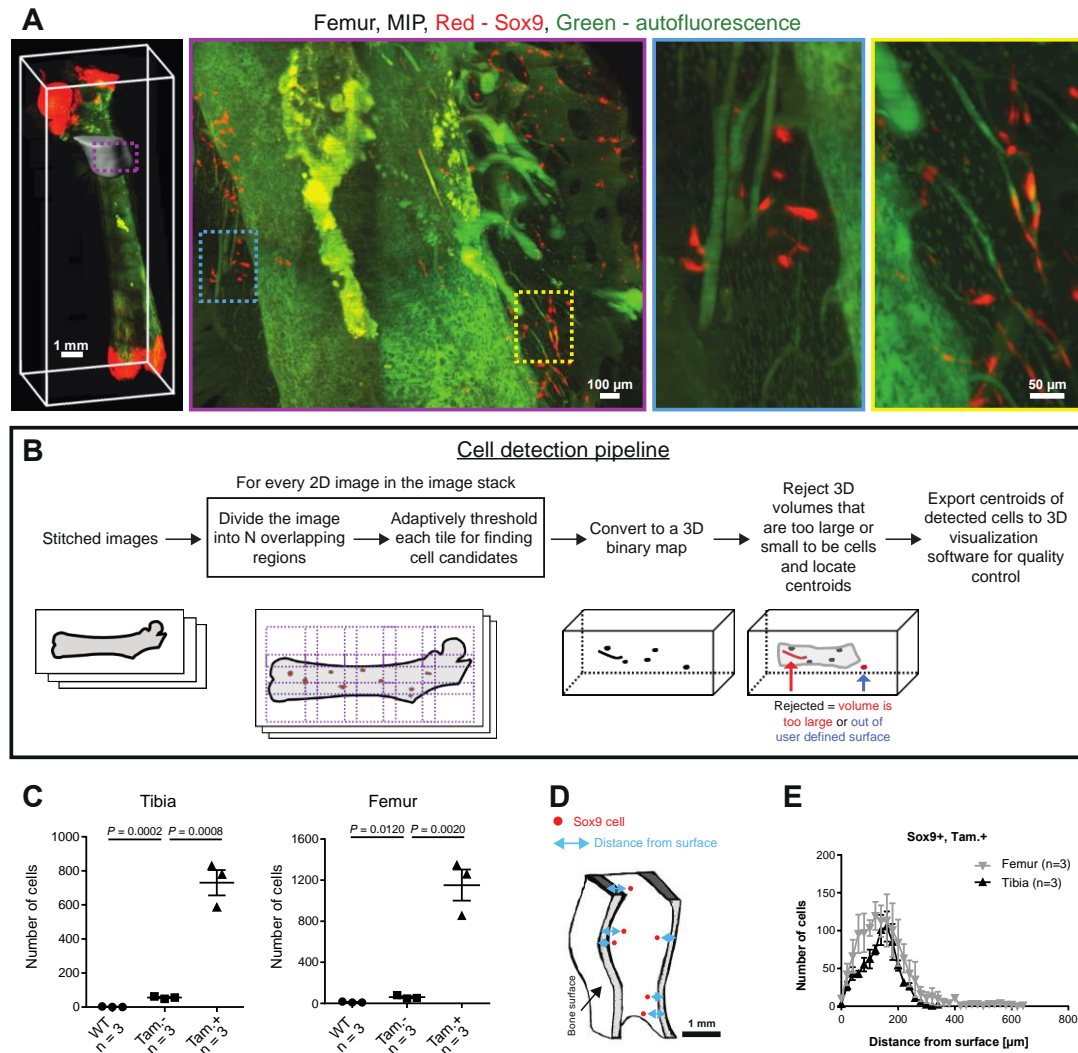
Other fields might also benefit from applying the bone clearing technique, such as neuroscience. Removal of the skull damages the interface between the skull and the underlying vascular bed and neuronal tissue. Preservation of this interface would be beneficial for studying the lymphatic vessels residing within it, assessing head trauma (e.g. percussive injuries), and characterizing the positioning of head-mounted brain implants in an intact environment<sup>9,63</sup>.

### 3.5 Main figures



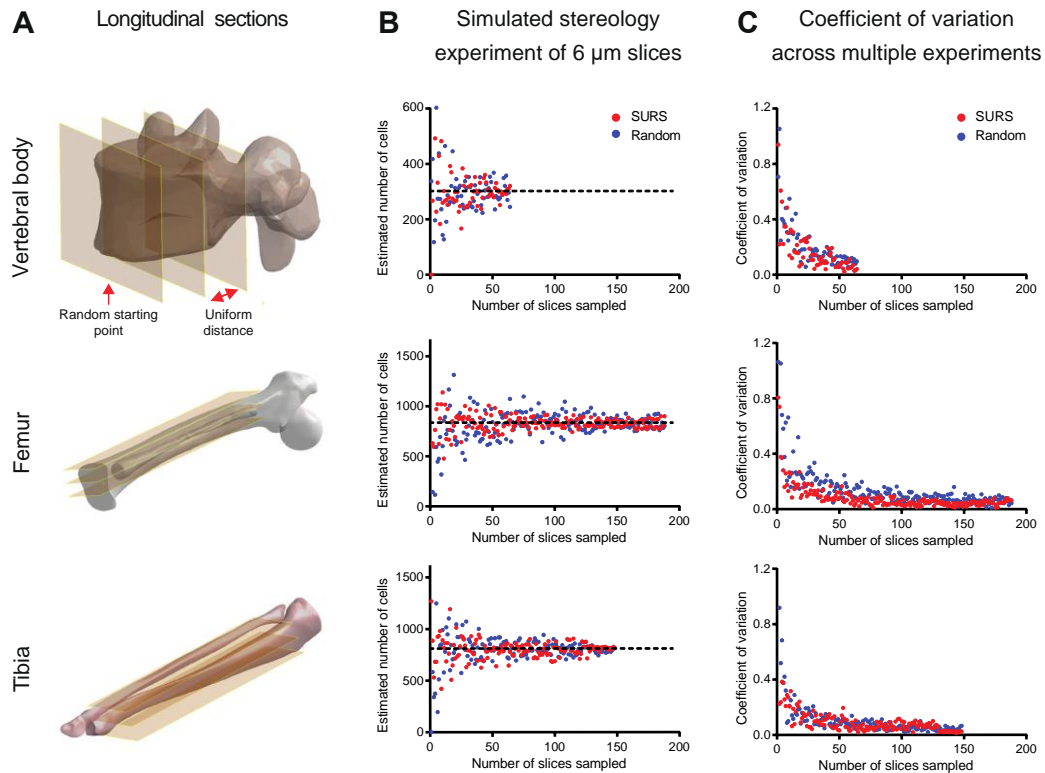
**Fig. 1. Bone CLARITY renders intact bones transparent while preserving endogenous fluorescence.** (A) Micrographs of mouse tibia, femur, and vertebral column before and after CLARITY. Bones were rendered transparent using deCAL CLARITY. (B) Block diagram outlining of the key steps of Bone CLARITY sample preparation. The procedure includes: demineralization, hydrogel stabilization, lipid removal via constant flow, and autofluorescence removal. (C) A schematic diagram of the imaging and computational

pipeline that utilizes the interface between 3D visualization and image-processing programs. **(D)** Maximum intensity projection (MIP) fluorescent image of the tibia and zoomed in digital slices (red = Sox9 and green = autofluorescence). Using Bone CLARITY, bones were imaged from one end to the other in the diaphysis (magnified images) and approximately 1.5 mm deep into the epiphysis (movies S1-3). Dotted yellow and blue boxed regions in the MIP represent the area of zoom in the digital slices (30  $\mu$ m thick). The purple and red dotted boxed regions in the digital sections represent the area that is shown with further higher magnification. The white dotted lines represent the boundaries of the compact bone. **(E)** MIP fluorescent image of the vertebral body.

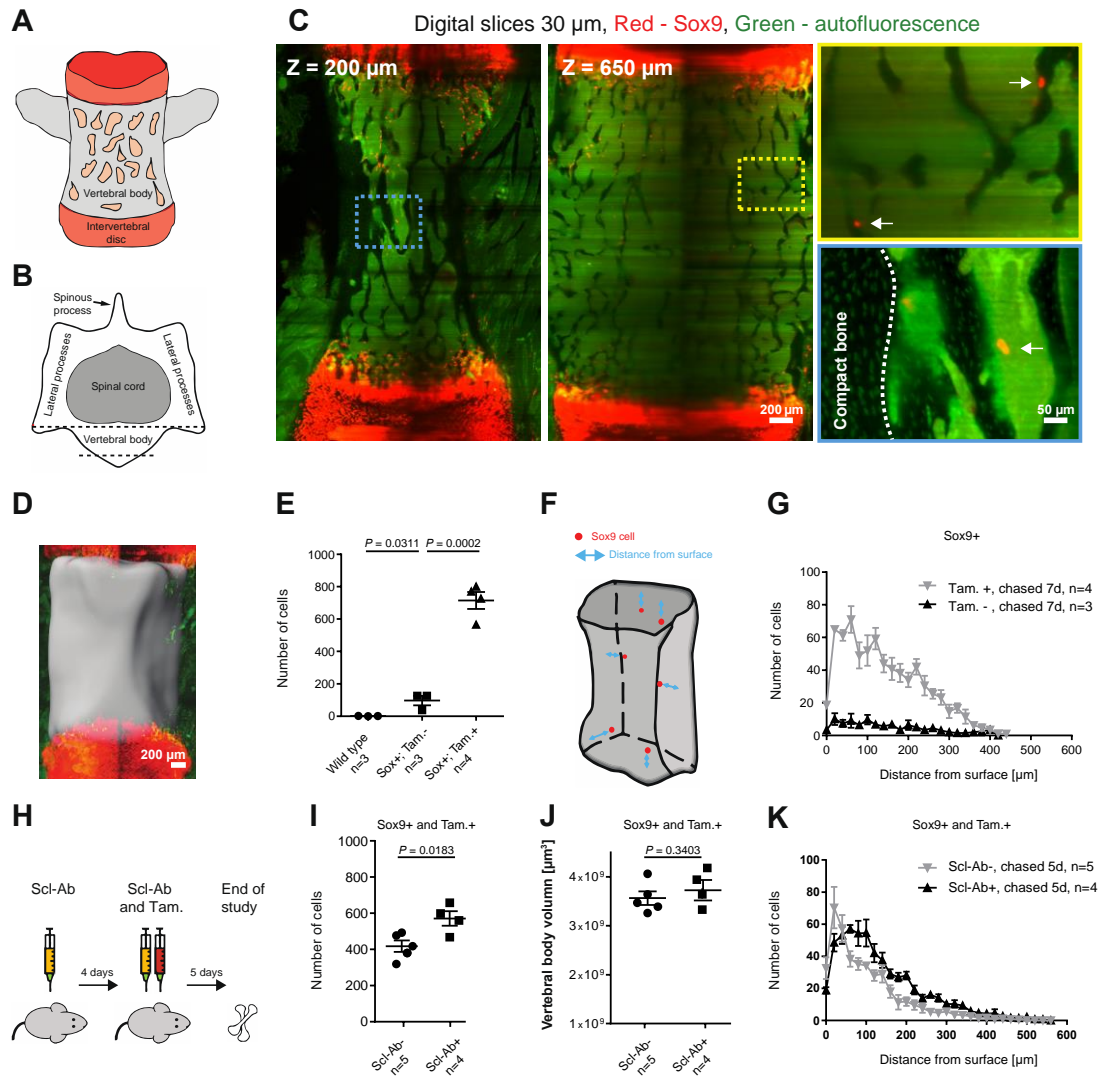


**Fig. 2. Bone CLARITY enables quantification of fluorescently labeled Sox9<sup>+</sup> cells in the mouse tibia and femur.** (A) Femur MIP fluorescent image and magnified images showing single Sox9<sup>+</sup> cells in the vicinity of the third trochanter (red = Sox9 and green = autofluorescence). The gray surface surrounding the femur represents an overlay of the volume of interest (VOI); only the cells that reside within the VOI are quantified. The purple, blue and yellow boxed regions in the MIP represent progressive magnification. (B) Block diagram of the semi-automated cell detection pipeline where cell candidates are identified via adaptive thresholding and their volume calculated. A predetermined selection criterion

based on cell volume is used to reject blood vessels or any detected blobs that are too large or too small. **(C)**. A comparison of fluorescently labeled Sox9<sup>+</sup> cell numbers between the experimental: tamoxifen positive (Tam<sup>+</sup>) and control groups: wild-type (WT) and tamoxifen negative (Tam<sup>-</sup>) for the femur and tibia; n = 3 animals per group; 1 bone per animal. For tibia, WT versus Sox9<sup>+</sup>;Tam<sup>-</sup>,  $P = 0.0002$  and Sox9<sup>+</sup>;Tam<sup>-</sup> versus Sox9<sup>+</sup>;Tam<sup>+</sup>,  $P = 0.0008$ . For femur, WT versus Sox9<sup>+</sup>;Tam<sup>-</sup>,  $P = 0.012$  and Sox9<sup>+</sup>;Tam<sup>-</sup> versus Sox9<sup>+</sup>;Tam<sup>+</sup>,  $P = 0.002$ . All values are mean  $\pm$  s.e.m; two-tailed, unpaired t-test. **(D)** Schematic illustrating how the distance of Sox9<sup>+</sup> cells from the surface is calculated. **(E)** The distribution of the number of Sox9<sup>+</sup> cells versus the distance from the bone surface.



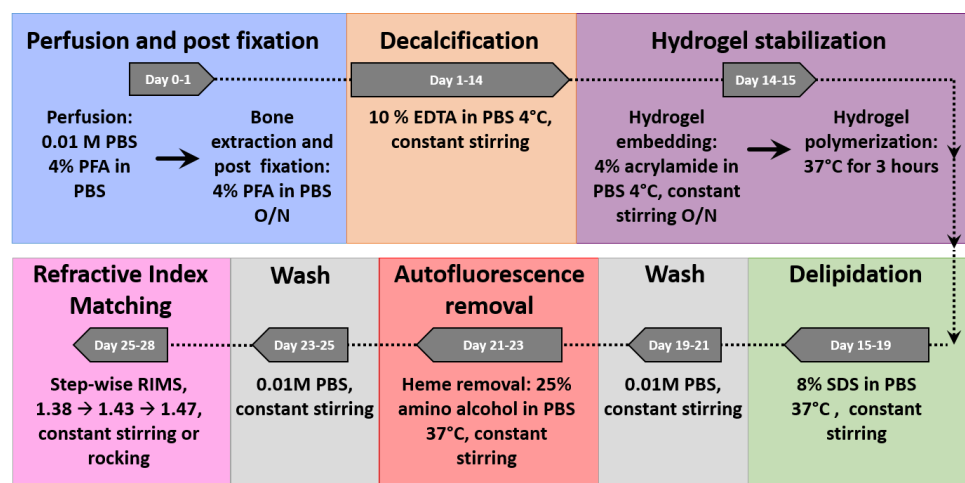
**Fig. 3. Using intact tissue clearing methods reduces cell estimate variability compared with traditional tissue sectioning methods.** (A) Schematic diagram showing the sample selection procedure for systematic uniform random sampling (SURS) of the mouse femur, tibia, and vertebral body.  $N$  uniformly spaced 2D sections are selected, offset by a random starting distance. The red single arrow indicates the random starting point, while the double arrow indicates the distance between the 2D sections. (B) Sox9<sup>+</sup> cell number estimation as a function of the number of sampled sections in a representative simulated stereology experiment using both SURS (red dots) and simple random sampling (blue dots). To simulate a stereology experiment, the cell density in the sampled sections is used to interpolate the cell number in the entire VOI. The black dashed line represents the ground truth, cell number based on the entire volume. (C) The coefficient of variation of five simulated stereology cell number estimates as a function of the number of sampled slices using SURS (red dots) and simple random sampling (blue dots).



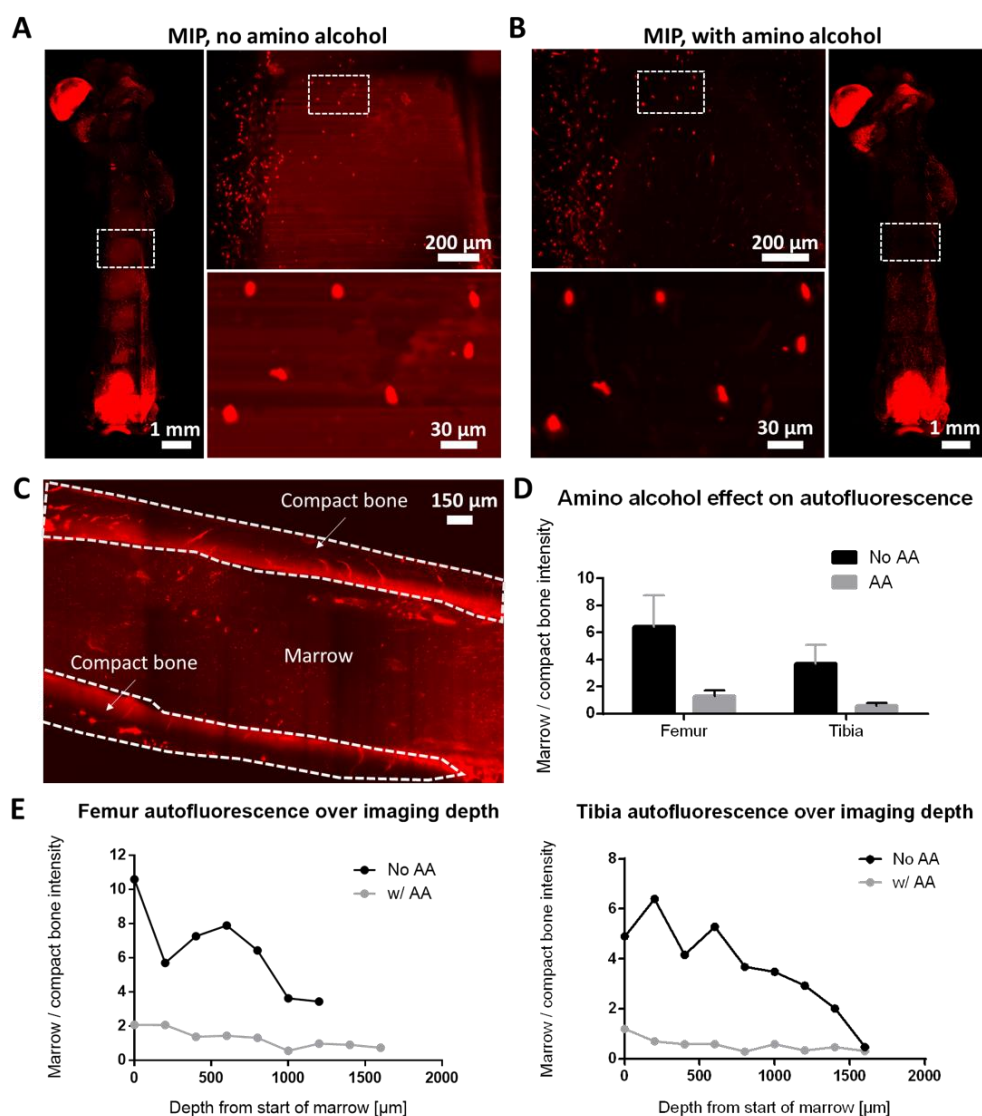
**Fig. 4. Bone CLARITY enables quantification of the effect of sclerostin antibody on the number of fluorescently labeled Sox9 cells in the mouse vertebra.** (A) Schematic depicting the lateral view of a mouse vertebra. The intervertebral disk is labeled in red; chondrocytes proliferate there after differentiation from tdTomato-Sox9<sup>+</sup> cells (chased for 7 days). The marrow is represented in light orange. (B) Transverse view of an L4 vertebra. The cell counts are isolated to the vertebral body, and the dashed lines approximate the locations of the sections shown in C. (C) Digital sections (30  $\mu\text{m}$  thick) at different depths along the vertebral body ( $Z = 200 \mu\text{m}$  and  $650 \mu\text{m}$ ), red = Sox9 and green = autofluorescence. The

intervertebral disks are notable at the vertical boundaries of the images. The white arrows in the color-coded magnified images point to Sox9<sup>+</sup> cells, and the dashed white line in the magnified image marks the interface between the compact bone and the marrow. **(D)** Micrograph of the mouse vertebral body showing a representative VOI (gray surface), which does not include the intervertebral disks. **(E)** Total cell number in tamoxifen-injected mice versus control groups. For WT versus Sox9<sup>+</sup>;Tam.<sup>-</sup>,  $p = 0.0311$  and Sox9<sup>+</sup>;Tam.<sup>-</sup> versus Sox9<sup>+</sup>;Tam.<sup>+</sup>,  $p = 0.0002$ . **(F)** Schematic illustrating how cell distance from surface is determined on a transparent vertebral body VOI. For each cell, the distance is defined as the shortest path to the surface in any direction. **(G)** The distribution of Sox9<sup>+</sup> cells as a function of distance from the endocortical surface. **(H)** Schematic of the sclerostin antibody (Scl-Ab) experimental timeline. **(I)** Quantification of the number of cells within the VOI outlined in **(D)** in the Scl-Ab and vehicle treated mice. For Sox9<sup>+</sup>;Tam.<sup>+</sup>;Scl-Ab<sup>-</sup> versus Sox9<sup>+</sup>;Tam.<sup>+</sup>;Scl-Ab<sup>+</sup>,  $p = 0.0183$ . **(J)** No significant change in the vertebral body volume was observed between the Scl-Ab group and the vehicle group. For Sox9<sup>+</sup>;Tam.<sup>+</sup>;Scl-Ab<sup>-</sup> versus Sox9<sup>+</sup>;Tam.<sup>+</sup>;Scl-Ab<sup>+</sup>,  $p = 0.3403$ . **(K)** The distribution of Sox9<sup>+</sup> cells as a function of distance from the endocortical surface. All values are mean  $\pm$  s.e.m; two-tailed, unpaired t-test.

### 3.6 Supplementary Figures

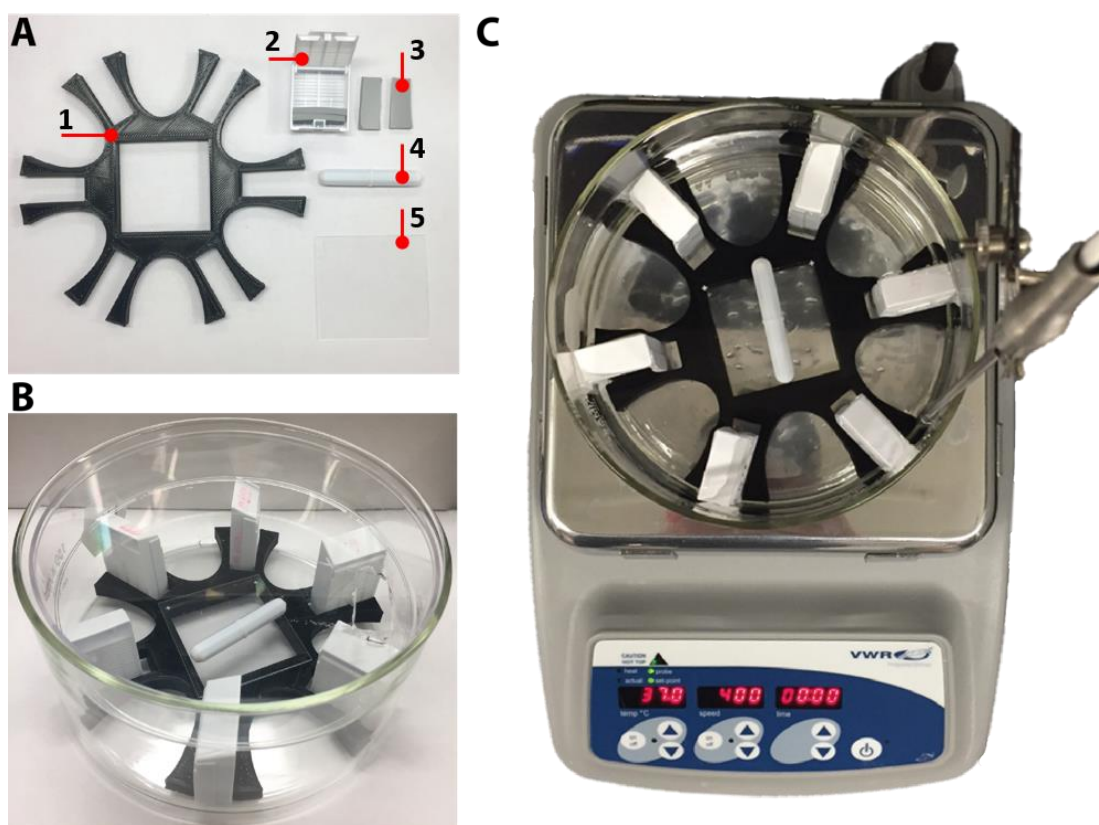


**Fig. S1. Bone CLARITY clearing process.** (A) As a first step (blue) in Bone CLARITY method transcathal perfusion is performed with 0.01M PBS, pH 7.4, and 4% paraformaldehyde (PFA) in PBS, pH 7.4. Hard tissue is then extracted and post-fixed with 4% PFA in PBS, pH 7.4 at 4°C for 16 hours. Hard tissue is then decalcified (orange). During this step, daily-exchange of 10% EDTA in PBS, pH 8, at 4°C under constant stirring is performed. Hydrogel stabilization to prevent protein loss is carried out (purple). The tissue is incubated in a hydrogel composed of 4% acrylamide with 0.25% VA044 in PBS at 4°C under constant stirring for 16 hours. Afterwards, the hydrogel is degassed with nitrogen gas and polymerized at 37°C. Delipidation is carried out with 8% SDS in PBS, pH 7.4 at 37°C under constant stirring (green). A wash step with PBS is performed before heme removal from the tissue (gray). Heme removal is performed with 25% amino alcohol in PBS, pH 9, at 37°C under constant stirring (red). A second wash step is performed on the tissue (gray). Finally, the tissue is refractive index matched to 1.47 through daily step-wise buffer exchange (pink).

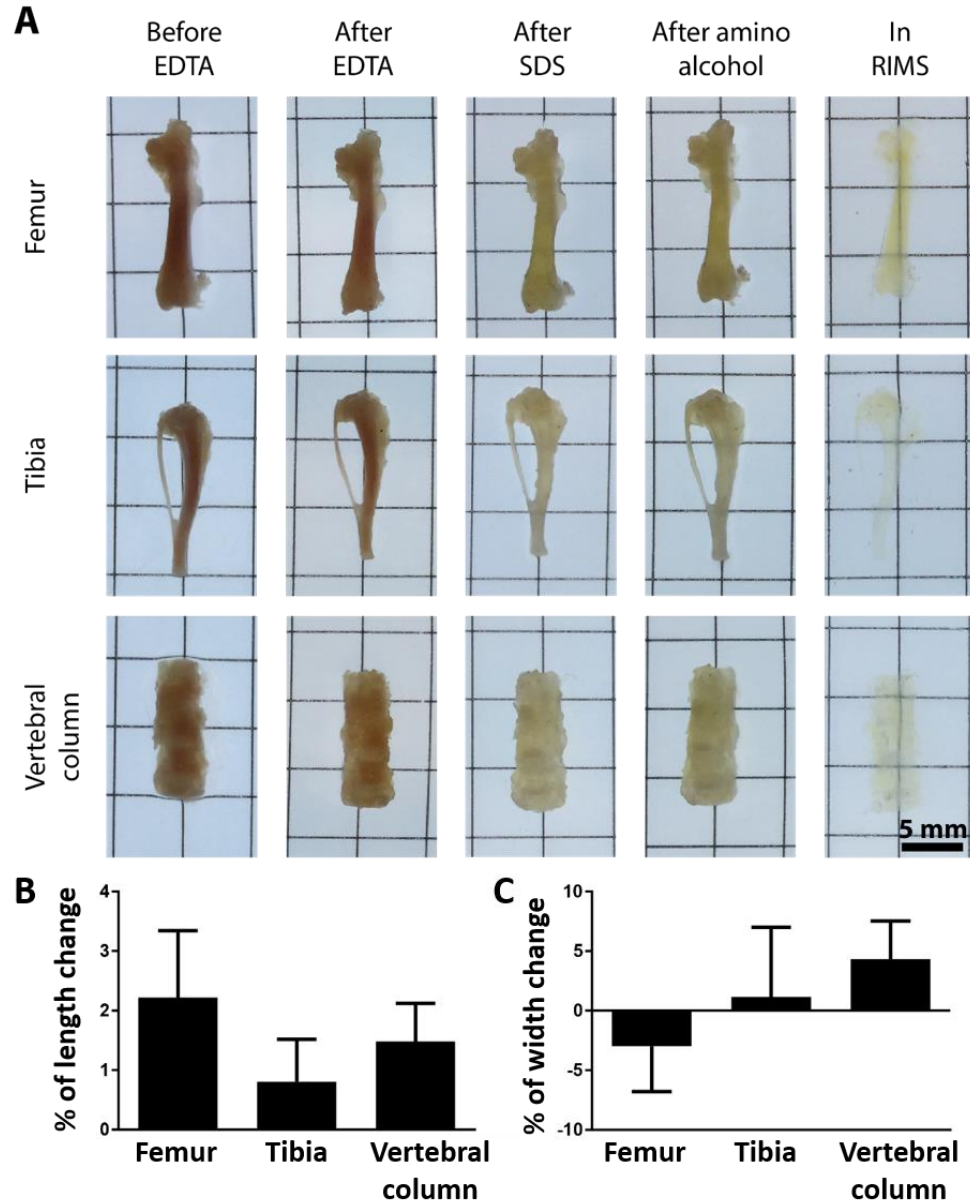


**Fig. S2. The effect of amino alcohol on reducing autofluorescence.** (A) MIP of a mouse femur rendered transparent using Bone CLARITY without the addition of amino alcohol. The dashed regions show the locations of progressively zoomed in images. (B) MIP of the same mouse femur shown in (A), after the addition of amino alcohol. Zoomed in images of the same regions as in (A) show that the same spatial pattern in the sample was maintained after amino alcohol treatment. (C) A representative image of the regions of tissue (marrow and compact bone) used for the analysis of the autofluorescence quenching effect. (D) The mean ratio of marrow to compact bone intensity in a mouse femur and tibia before and after

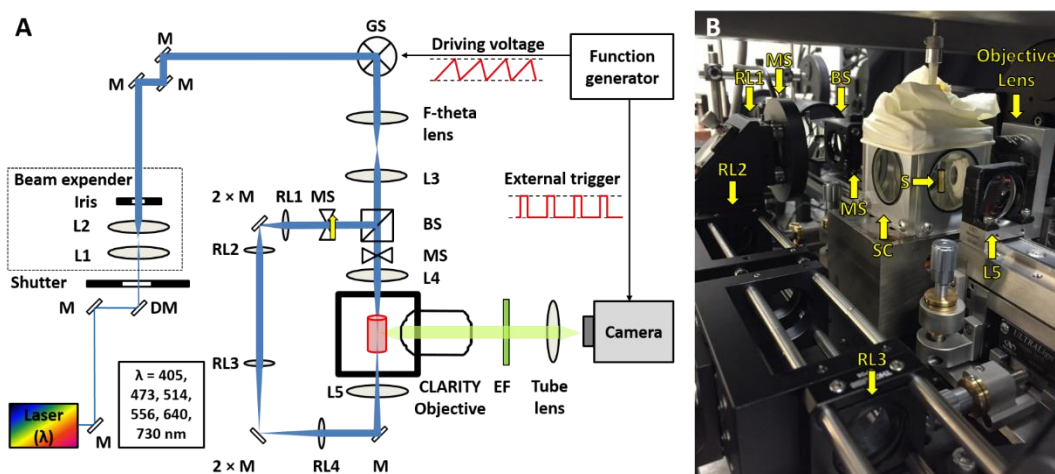
the addition of amino alcohol - which removes autofluorescence in the bone marrow. The intensity ratio between the marrow and compact bone should be independent of the imaging depth and illumination variations. **(E)** The ratio of marrow to compact bone intensity from different sample regions along the imaging depth of the bone. A depth of 0  $\mu\text{m}$  refers to the first region along the imaging depth that had both marrow and compact bone.



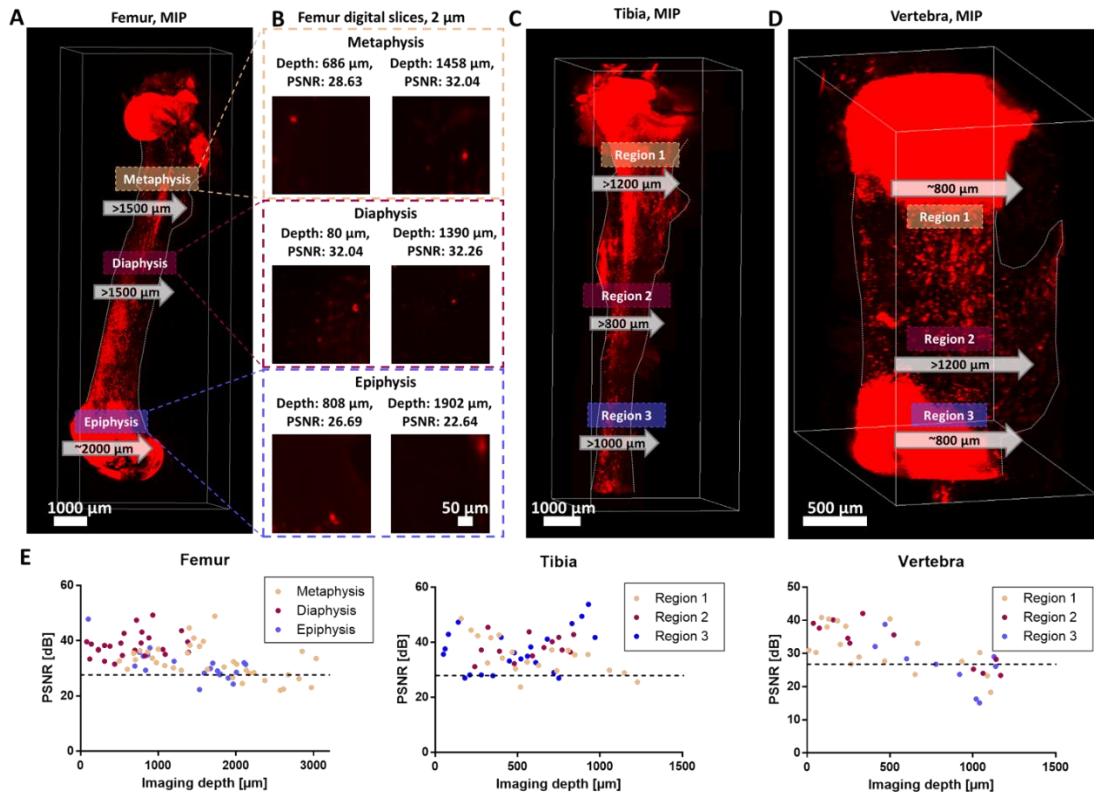
**Fig. S3. Bone CLARITY setup.** (A) The individual parts of the setup, which is composed of: (1) A 3D printed holder for histology cassettes. (2) A large histology cassette with rubber inserts. (3) Rubber inserts to prop the tissue above the 3D printed holder, and consequently position the tissue within the convective current. (4) A stir bar. (5) A glass slide for placement of the stir bar above the 3D printed holder. (B) The histology cassettes arranged inside a 15 cm petri dish with a 14.7 cm plastic petri dish cover. (C) The petri dish is placed on a closed loop temperature controlled stir plate. For repeatability, six cassettes are always placed in the chamber, and the stir bar is rotated at a speed of 400 rotations per minute at 37°C. The design file for the 3D printed holder is available as an auxiliary supplementary file.



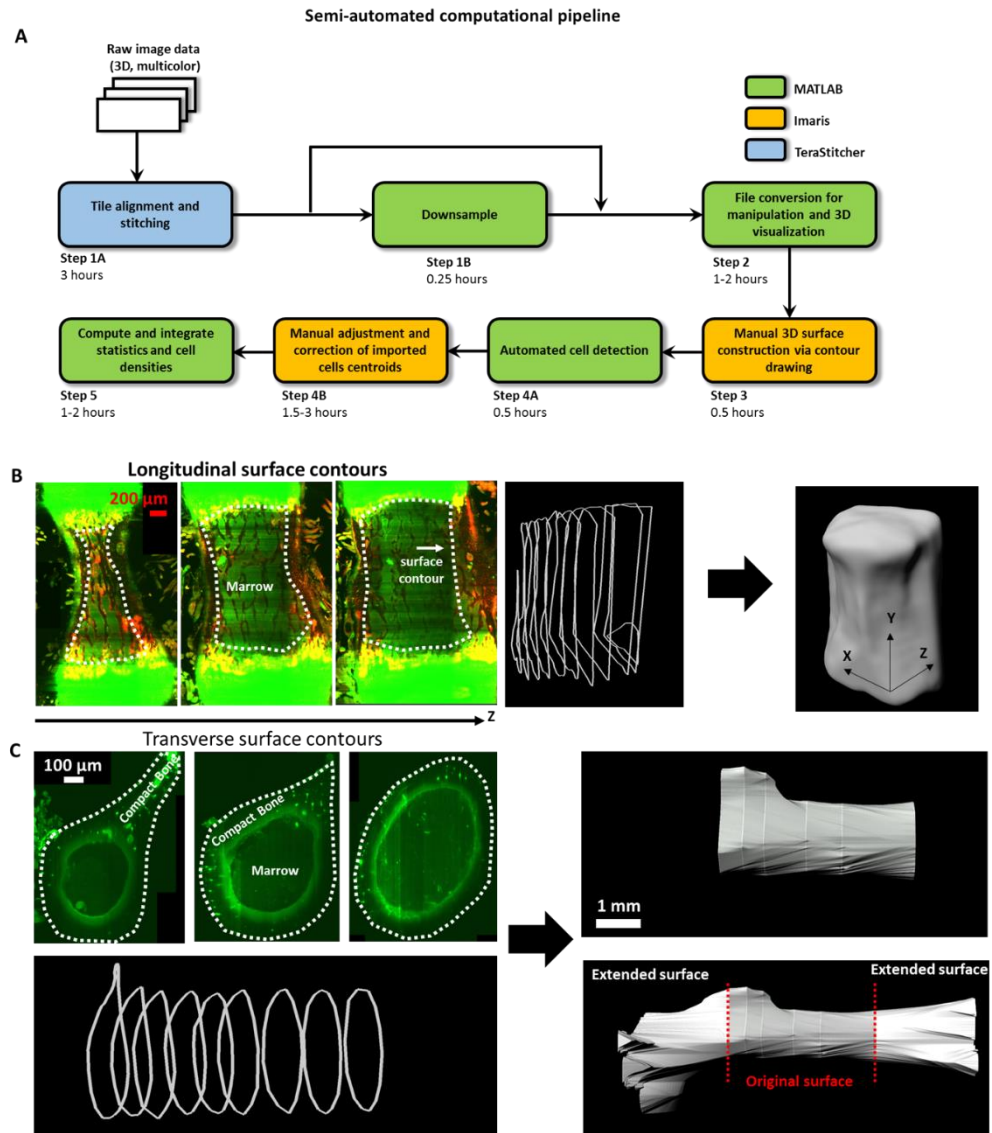
**Fig. S4. Sample progression through Bone CLARITY clearing process.** (A) Micrographs documenting the progression of the femur, tibia and vertebral column throughout each phase of Bone CLARITY. (B) The percent change in length of the bones before and after Bone CLARITY. (C) The percent change in width of the bones before and after Bone CLARITY. For (B) and (C),  $n = 6$  for the femur,  $n = 5$  for the tibia and  $n = 4$  for the vertebral column. All values are mean  $\pm$  s.e.m.



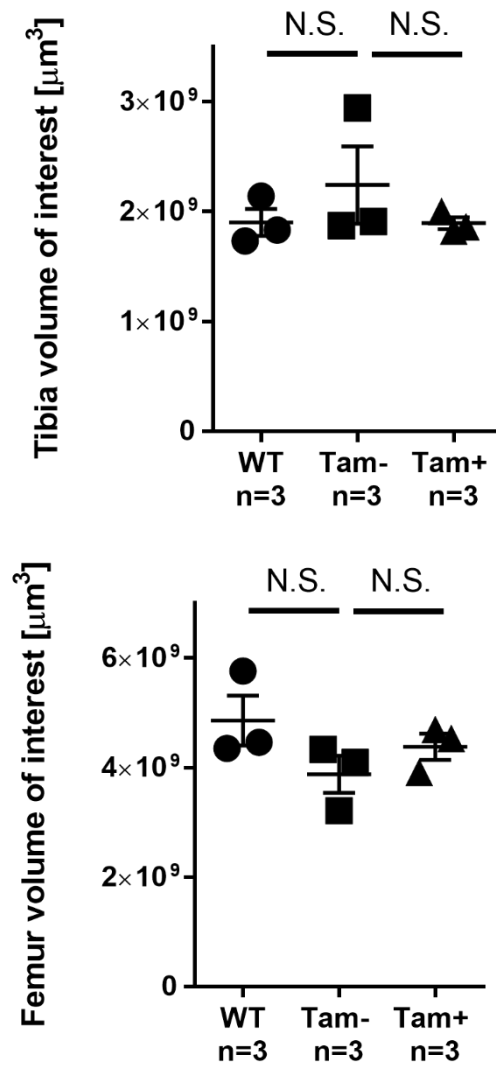
**Fig. S5. LSFM set-up.** (A) Schematic diagram of the multi-color dual-side illumination light-sheet microscope with the confocal slit detection. The microscope acquisition speed is up to 22 frames per second; consequently, it allows for high-resolution acquisition of large cleared tissue. L – lens, M – mirror, DM - dichroic mirror, GS – galvano scanner, BS – beam splitter, MS – mechanical shutter, RL – relay lenses and EF emission filter. (B) A photograph of the light-sheet setup, the sample (S) is directly immersed in the sample chamber (SC).



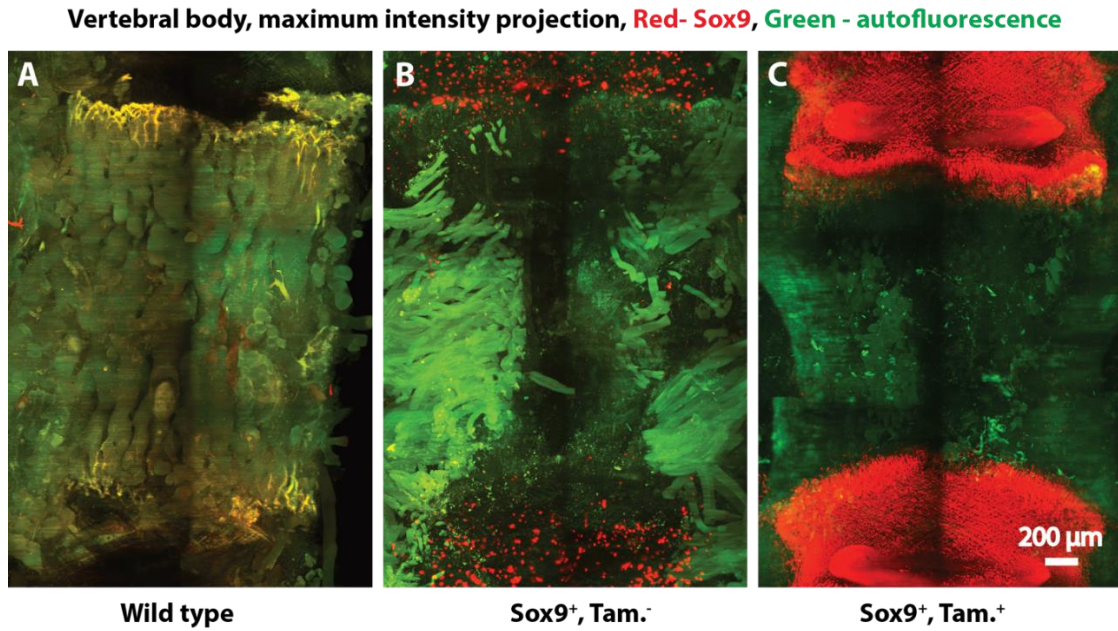
**Fig. S6. Signal quality metrics to quantify imaging depth of Bone CLARITY.** (A) MIP of a cleared mouse femur with estimates of imaging depth limit at different anatomical regions. The arrows indicate the direction of analysis and the calculated imaging depth. (B) Examples of isolated Sox9<sup>+</sup> cells at different imaging depths and in different bone regions. Sox9<sup>+</sup> cells remain distinguishable at the deepest parts of the metaphysis and diaphysis, but signal loss is evident in the deeper parts of the epiphysis. The same analysis was repeated for a cleared mouse tibia (C) and vertebra (D). (E) The peak signal-to-noise ratio (PSNR) of isolated Sox9<sup>+</sup> cells. The dashed lines represent a qualitative estimate of where cells are no longer distinguishable from background noise, at a PSNR of around 24.



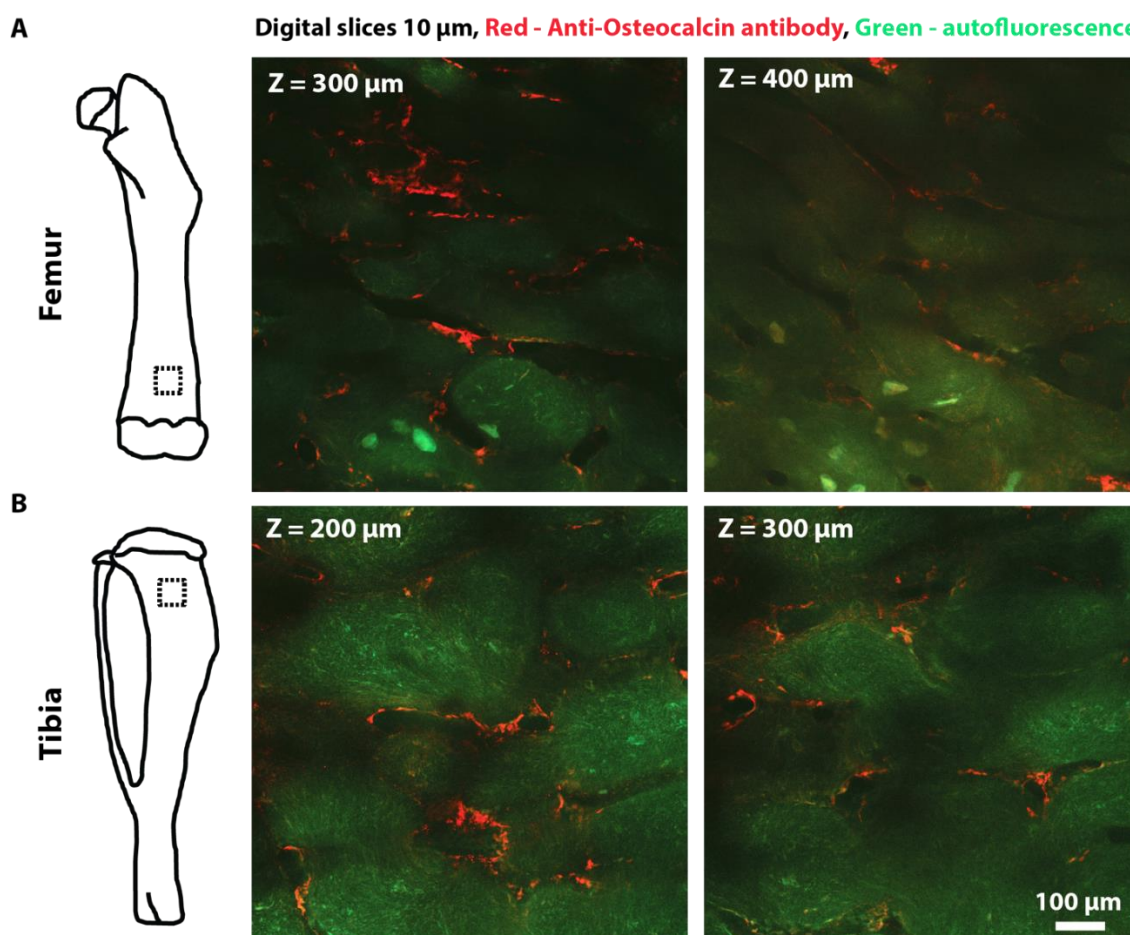
**Fig. S7. Computational pipeline.** (A) A block diagram of the computational pipeline. (B) VOI creation in the vertebra body. In each block of approximately 10 longitudinal sections a ROI (dotted line) is marked around regions that contain marrow. The VOI is then interpolated from the ROIs. (C) VOI creation in the tibia and femur, where approximately 10 transverse slices are used to create the VOI. For cell distance to surface calculation, the VOI is extended from both ends, in order to ensure distance calculation to the real bone boundaries and not to artificial perpendicular surfaces.



**Fig. S8. Volume comparisons in the tibia and femur.** No significant change in the volume of interest was observed between the Sox9<sup>+</sup> and Tam<sup>+</sup> group and the two control groups (wild type; Sox9<sup>+</sup> and Tam<sup>-</sup>). For all groups n = 3, all values are mean  $\pm$  s.e.m; two-tailed, unpaired t-test, ( $p \geq 0.05$  = N.S.).



**Fig. S9. TdTomato expression in the vertebral columns.** (A) No fluorescence signal is present in the emission spectrum of tdTomato (556 nm laser excitation) in the vertebral column of a wild-type mouse. (B) Using the same imaging settings as in panel (A), Sox9<sup>+</sup> mouse shows tdTomato expression even without tamoxifen induction. The fluorescence signal is primarily located at the endplates and attributed to leakage. (C) Using the same imaging settings as panel (A), Sox9<sup>+</sup> mouse shows high levels of tdTomato expression after tamoxifen induction.



**Fig. S10. Antibody staining using Bone CLARITY.** (A) The approximate region imaged in the bisected femur is outlined by a black square. Digital slices of 10  $\mu\text{m}$  thickness are shown in 300 and 400  $\mu\text{m}$  from the surface of the femur. Red shows Osteocalcin staining of cells lining the surface of bone, green shows autofluorescence. (B) The region imaged in the bisected tibia is outlined by a black square. Digital slices of 10  $\mu\text{m}$  thickness are shown in 200 and 300  $\mu\text{m}$  from the surface of the tibia. Red shows Osteocalcin staining of cells lining the surface of bone, green shows autofluorescence.

### 3.7 Supplementary movie captions

**Movie S1: Visualizing endogenous fluorescence throughout a cleared mouse femur.**

The movie shows a MIP image of an entire mouse femur at different angles and zoomed in view of the greater trochanter area. Digital sections with 30  $\mu\text{m}$  thickness are shown at different bone depths at the greater trochanter. Sox9<sup>+</sup> cells are labeled with red fluorescent protein (tdTomato), while tissue autofluorescence in green provides structural cues. (doi:[10.22002/D1.234](https://doi.org/10.22002/D1.234))<sup>1</sup>

**Movie S2: Visualizing endogenous fluorescence throughout a cleared mouse tibia.**

MIP image of an entire mouse tibia rotated in different angles. Zoomed in digital sections (30  $\mu\text{m}$  thick) of the tuberosity and the diaphysis are shown. The Sox9<sup>+</sup> cells are labeled with tdTomato and tissue autofluorescence in green for structural orientation. When the red channel is turned off ( $Z = 2016 \mu\text{m}$ ) some cellular processes can be seen. (doi:[10.22002/D1.235](https://doi.org/10.22002/D1.235))<sup>1</sup>

**Movie S3: Visualizing endogenous fluorescence throughout a cleared mouse vertebral**

**body.** The movie shows an L4 vertebra body of a mouse. The intervertebral disks labeled in red can be seen in the vertical edge of the image. Longitudinal sections (30  $\mu\text{m}$  thick) of the vertebra body at different depths can also be seen where the red channel shows Sox9<sup>+</sup> cells and the green channel shows tissue autofluorescence. (doi: [10.22002/D1.236](https://doi.org/10.22002/D1.236))<sup>1</sup>

<sup>1</sup> = data file is stored at [data.caltech.edu](https://data.caltech.edu)

### 3.8 Materials and methods

**Study design:** The objective of this study was to enable the visualization and quantification of cell population in an intact bone tissue by developing and integrating tissue clearing, fluorescence microscopy, and a computation pipeline. Experimental and control animal cohorts were chosen based on preliminary data that suggested a large effect size. All transgenic animals used in this study are as described in the animals section. All wild-type animals used in this study were C57BL/6. To characterize the density of sox9<sup>+</sup> cells and distribution within the femur, tibia and vertebral column, male transgenic animals of 6-7 weeks of age received a 2 mg tamoxifen IP on day 1 of the experiment to enable expression of a native fluorescent gene for 7 days before culling. For the study of the effects of sclerostin-antibody on total number of osteoprogenitor cells, a cohort of 7 weeks old male transgenic animals were treated with a sclerostin antibody at 100 mg/kg IP on day 1, and again 4 days later at 100 mg/kg IP with tamoxifen induction at 2 mg IP before culling on day 9 of the study. The sclerostin antibody was provided by Amgen. During cell counting all manual quantification is performed in a blind manner to eliminate observer bias. Animals were randomly assigned to groups for experiments. Raw data values for cell counts are reported in an auxiliary supplementary file.

**Animals:** A transgenic Sox9-creERT2 mouse (RIKEN BioResource Center, #RBRC05522) was crossed to a Rosa26-loxP-stop-loxP-tdTomato (R26R-tomato, JAX7914) reporter mouse to generate Sox9-CreERT2-R26-tdt. Sox9-creERT2-R26-tdt mice ages 6-7 weeks old were used in the experiments involved in labeling Sox9 cells (n = 18). The labeling was achieved by the injection of 2 mg tamoxifen IP, which was dissolved first in 100% ethanol then in sunflower seed oil (Sigma, #S5007) overnight at 60°C. Mice were genotyped by PCR. All mice were analyzed in mixed backgrounds. Mice were group housed in sterile, ventilated micro-isolator cages on corn cob bedding in an AAALAC accredited facility. All procedures were conducted in compliance with the Guidelines for the Care and Use of Laboratory Animals approved by Massachusetts General Hospital's Institutional Animal Care and Use Committee (IACUC). Animals were provided ad libitum access to pelleted feed (LabDiet 5010) and water (Standard drinking water of Boston,

Massachusetts, pH 7.8) via HYDROPAC. Animals were maintained on a 12:12 hour light:dark cycle in rooms at 64-79 F with 30-70% humidity under pathogen-free conditions.

***Bone deCAL CLARITY protocol:*** The clearing process is summarized in fig. S1. After euthanization, mice were perfused transcardially with 0.01M PBS (Sigma, #P3813) followed by 4% paraformaldehyde (PFA) (VWR, #100496-496) and the femurs, tibias and L3-5 vertebral columns were extracted. The bones were post-fixed overnight in 4% PFA. To enhance clearing of hard-tissue, the demineralization phase was extended to 2 weeks with 10 % EDTA (Lonza, # 51234) in 0.01M PBS, pH 8. During the demineralization phase, samples were kept under constant stirring in histology cassettes (Electron Microscopy Sciences, #70077-W) at 4°C with fresh EDTA buffer exchanges daily. Next, the decalcified bones were embedded in a hydrogel matrix (A4P0) which consist of 4% acrylamide (Bio-Rad, #1610140), 0% paraformaldehyde, and 0.25% thermo-initiator A4P0 (Wako Chemicals, VA-044) in 0.01M PBS overnight at 4°C. The samples were degassed through nitrogen gas exchange for 5 minutes and polymerized at 37°C for 3 hours. After structural reinforcement with the A4P0 hydrogel, delipidation was carried out with 8% SDS in 0.01M PBS, pH 7.4, for 4 or 5 days (vertebral body and long bones respectively) at 37°C under constant stirring (fig. S3). The samples were then washed for 48 hours in 0.01M PBS with 3 buffer replacements. The amino alcohol N,N,N',N'-tetrakis(2-hydroxypropyl)ethylenediamine (Sigma, #122262-1L) was added at 25% W/V in 0.01M PBS, pH 9, for 2 days at 37 °C under constant stirring for the purpose of decolorization of the tissue through heme group removal. Lastly, the bones were washed with 0.01M PBS for 24 hours and subsequently immersed in refractive index matching solution (RIMS). The bones were gradually immersed in RIMS (15) with a refractive index of 1.47 through daily step-wise RIMS exchange starting with RIMS 1.38, RIMS 1.43 and finally RIMS 1.47.

***Light-sheet microscopy imaging:*** Prior to imaging, the sample was placed in the LSFM sealed immersion chamber for at least three hours, allowing the RIMS solution in the chamber to equilibrate with the residue RIMS in the bone sample. To minimize optical aberrations, the refractive index of the RIMS solution was measured and the correction

collar on the objective lens (10× CLARITY Objective lens with numerical aperture of 0.6, Olympus XLPLN10XSVM) was set accordingly. To image the entire bone, multiple tiles with 10% overlap were acquired. Typically, the femur, tibia and vertebral body required 13×5, 11×5 and 3×2 tiles respectively (vertical × horizontal). In a calibration stage that took place prior to the scan, the following parameters were defined for each tile: (i) Light-sheet illumination direction; the LSM has two light-sheets that illuminate the sample from opposite directions. Selecting the preferable illumination direction dramatically reduced scattering. (ii) The start and end point of the Z-stack; this step was done in order to minimize the number of acquired images in an already big data set (50-500 GB). (iii) The focus points of the detection objective along the scan were defined to mitigate RI variations along the scan that created out-of-focus aberrations. Once the calibration stage was completed, the bone was imaged with a frame rate of 22 frames per second, and bit depth of 16 bits. The acquired data set size depends on the sampled voxel size. For a voxel size of  $0.585 \times 0.585 \times 2 \mu\text{m}^3$  the tibia and femur produced ~ 250 GB of data per color channel, while the vertebral body produced ~ 30 GB of data per color. Generally, the datasets are down-sampled after acquisition for processing; the typical voxel sizes are  $1.17 \times 1.17 \times 2 \mu\text{m}^3$  and  $2.34 \times 2.34 \times 2 \mu\text{m}^3$  for the vertebra column and long bones respectively.

All experimental and control groups were imaged with the same laser power. For images that were acquired deep in the bone and when the SNR changed within the distance from the bone boundary, the contrast and gamma were adjusted in the displayed images. The gamma adjustment was done to visualize cells that exhibit both low and high intensity within the same field-of-view. Images from the vertebra (Sox9<sup>+</sup> and Tam<sup>+</sup> group) are representative of 13 vertebrae from 13 mice. Images from the tibia and femur (Sox9<sup>+</sup> and Tam<sup>+</sup> group) are representative of 5 tibias and 5 femurs from 5 mice.

**Statistical analysis:** All statistical analysis was performed using GraphPad Prism version 7.01. For figures 2C, 4E, 4I and 4J, mean values for each group were compared using an unpaired t-test. In all graphs, data points per individual animal trial with the mean value and s.e.m are shown.

### 3.9 Supplementary methods

*Computational pipeline:* LSFM generates multidimensional datasets on the order of hundreds of gigabytes, for example 500 GB for tibia with 2 color channels, motivating a necessity for high-throughput automated computational tools to do data processing, management, and visualization. Thus, we established a computational pipeline optimized for large datasets that interfaces between 3D visualization and analysis software (Imaris, Bitplane) and a programming framework (Matlab, Mathworks) although alternatives can also be used (e.g. Amira, Python).

Tile stitching was done with TeraStitcher. The acquired tiles (.tiff format) were saved according to the TeraStitcher two-level hierarchy and file naming convention. The resulting stitched tiles were saved as large individual.tiff format file per Z section. To accelerate processing and manipulation of 3D images in Imaris, the lateral resolution was down-sampled by a factor of 2 and 4 for vertebrae and long bones, respectively. These image stacks were then converted into a single large multiresolution 3D Imaris image. The 3D image is then used to construct the VOI, performing quality control for the automatic cell detector, and obtaining relevant statistics about the position of the cells within the VOI.

Using Imaris surface tools, a VOI was constructed around the bone surface, while excluding connective tissue. The VOI was used to count cells and evaluate morphology (fig. S7). For the femur and tibia, the VOI was defined around known anatomical regions: the third trochanter and tibial crest, respectively. The surface for the vertebral body aimed to surround the inner marrow of interest and exclude surrounding compact bone and chondrocyte-dense endplates.

2D stitched images were used for automatic Sox9<sup>+</sup> cell detection (Fig. 2B). The algorithm iterates through the image stack and divides the 2D stitched images into a number of smaller regions. For each region, the standard deviation and mean value of fluorescence intensity is calculated. Sox9<sup>+</sup> cells are expected to be significantly brighter in intensity than the surrounding tissue, therefore each sub region is thresholded with a value of approximately 9 standard deviations above the mean value. The thresholded images are then converted to a

binary format and morphological operations are conducted to improve connectivity and discard noise. A 3D binary matrix is then constructed from the 2D binary images and connected voxels are grouped into 3D blobs. Blobs that showed one of the following criteria were removed: (i) blobs with volume outside of a  $523\text{--}33,500\ \mu\text{m}^3$  range, (ii) blob with spatial position outside of the VOI and (iii) blobs that show strong signal in the autofluorescence channel. The centroids of the remaining blobs are imported into Imaris to perform manual adjustments. The goal of the automatic detection is to facilitate manual detection by saving time and reducing the chance of error from long annotation sessions.

Using Imaris' clipping plane tool, the 3D image is divided into thick slices ( $\sim 100\ \mu\text{m}$ ) to allow for manual adjustments of the automatically detected cells. Viewing cells in these thick sections allows for visualization of 3D features that help to accurately identify cells. When counting the cells, 2 color channels were used to disambiguate between autofluorescence (green) and the true Sox9<sup>+</sup> signal (red). If the Sox9<sup>+</sup> cell under question is present with strong intensity in the autofluorescence channel, then it was not included in the quantification.

Following manual cell count adjustments, Imaris was used to generate distance to surface and cell count statistics. The surface was imported into Imaris' cell object with the marked Sox9<sup>+</sup> spots as its vesicles. Automatically generated statistics were imported to Matlab for further processing and plotting.

*Sample mounting for light-sheet microscope:* To mount the sample for LSFM, one end of a 4 cm long plastic capillary was flattened using a plier, and a small amount of All Purpose Krazy Glue (KG585) was applied on this end. The flattened end was later placed on the edge of the cleared condyles (femur) and L3 (vertebral column) bone while moderate pressure was applied. Four minutes later, the capillary bonded with the bone and the sample was placed back into RIMS. Before imaging, the plastic capillary was fitted to a glass capillary, which can be attached to the sample stage of the LSFM (fig. S5).

*Simulated stereology:* To investigate the variability between 3D counting and 2D section counting, we performed a series of simulated stereology experiments using our 3D

datasets. To start, the coordinates of each cell center were imported to MATLAB as a list, and the VOI was imported to MATLAB as a binary mask. From the binary mask, the volume of each 2D section was calculated by accumulating the number of pixels in the 2D section and multiplying it by the voxel size. Then, a simple random sampling stereology experiment was simulated by stochastically selecting  $N$  digital sections (6  $\mu\text{m}$  thick) from the entire VOI. Only cells whose centroids were within the volume of the selected digital sections were counted, this guarantees that each cell will be counted only once and cell will not be overrepresented. To calculate the cell density, the total number of counted cells in the selected digital sections was divided by their accumulated volume. From sampled cell density and the known volume of the entire bone, the total number of cells in the bone was estimated. In order to calculate the coefficient of variation, each experiment was repeated 5 times and the standard deviation of the experiments was divided by the total cell number of the volume. To reflect modern stereology sampling techniques, we also simulated systematic uniform random sampling (SURS) stereology experiments. As compared to simple random sampling, SURS has a stricter subset of possible samplings. Rather than looking at any random subset of slices, SURS requires sampling the tissue with a fixed interval, while the first starting slice is randomly selected (Fig. 3A). In order to compare SURS with simple random sampling, each SURS experiment consisted of a specific number of digital slices. Since a specific number of digital slices can be attained by a variety of sample intervals, the minimum interval that would yield the desired number of slices was selected. Once the digital sections were selected for SURS, the cell density was calculated in the same manner as simple random sampling. To calculate the coefficient of variation, each SURS experiment was repeated for 5 random starting positions. Figure 3C shows only numbers of slices that yielded at least 5 possible starting positions.

*Quantifying the effect of amino alcohol (related to fig. S2):* High levels of autofluorescence can overshadow elements of interest and lead to false inclusion of features. Intensity data from the mouse femur and tibia before and after treatment with amino alcohol (N,N,N',N'-tetrakis(2-hydroxypropyl) ethylenediamine) (Sigma, #122262-1L) were analyzed. The starting depth (0  $\mu\text{m}$ ) was defined as the first section that contained

both sufficient compact bone and marrow for analysis. Then, for 7-9 depths along the imaging dimension, two areas ( $\sim 1000 \mu\text{m}^2$ ) were extracted. The first area exclusively contained compact bone, while the other exclusively contained marrow. The mean intensities of the two areas (marrow and compact bone) were then calculated. The mean intensity ratio (marrow versus compact bone) was used to quantify the amino alcohol effect on quenching auto-fluorescence in the marrow, since amino alcohol operates on heme, which is found only in the marrow. To verify that amino alcohol did not quench endogenous fluorescence, we compared the same region of the bone before and after amino alcohol treatment (fig. S2).

*Quantifying imaging depth at different anatomical regions of the bone (related to fig. S6):* Bone tissue is difficult to clear and image in 3D, particularly in deeper regions where the SNR is low. Additionally, bones are highly heterogeneous and different regions have significantly different biological makeup that affects imaging SNR. In order to characterize the imaging depth capabilities of LSM in bone tissue cleared by Bone CLARITY, we quantitatively assessed cell detection in several anatomically distinct regions using a Peak SNR (PSNR) metric (equation 1).

$$PSNR = 20 \times \log_{10} \frac{MAX - \mu}{\sigma} \quad (1)$$

MAX is the maximum intensity of pixels in the region containing the cell,  $\mu$  is the mean intensity of all pixels in the region that is at least  $20 \mu\text{m}$  away from the cell center, and  $\sigma$  is their standard deviation. For each anatomical region, the starting point ( $0 \mu\text{m}$ ) of the sample was defined as the first depth that bone tissue filled at least 50% of the field-of-view. Then, at least 10 isolated Sox9<sup>+</sup> cells were selected along the image stack and their PSNR were calculated. For PSNR calculation, an area of  $0.123 \text{ mm}^2$  and  $0.086 \text{ mm}^2$  (for long bones and vertebrae, respectively) was used for the region around the cell. Based on visual inspection, the minimum PSNR for cell detection was 24. Consequently, we could reliably image bone tissue to a depth of approximately 1.5 mm. With the exception of chondrocyte-dense regions such as the growth plate and the epiphysis, this demonstrates that Bone CLARITY and LSM provide end-to-end imaging of mouse bone tissue. Note that in regions

with high concentration of chondrocytes, no isolated cells could be found, thus creating large gaps along the imaging depth in fig. S6E.

*Antibody staining:* Mice were perfused transcardially with 0.01M PBS (Sigma, #P3813) followed by 4% paraformaldehyde (PFA) (VWR, #100496-496). The femurs and tibias were extracted and post-fixed overnight with 4% PFA. The bones were demineralized with 10% EDTA in 0.01M PBS (Lonza, #51234) for two weeks at 4 degrees with daily solution exchange. Afterwards, the bones were incubated in 30% sucrose (Sigma, #S0389) in 0.01M PBS until they equilibrated to the solution. Next, the bones were embedded in OCT compound (Sakura Finetek, #M71484) in a tissue mold cassette and slowly frozen in powdered dry ice containing ethanol (Sigma, #E7023). The bones were then cut along its longitudinal plane with a cryostat (Leica, Biosystems) until one-half remained. The OCT compound was removed from the bisected bone and cleared following the Bone CLARITY protocol starting with hydrogel embedding and ending before refractive index matching stage. Antibody staining was performed using an anti-osteocalcin antibody (Abcam, #ab93876) at 1:100 in 0.5mM SDS with 0.01 % sodium azide (Ricca Chemical Company, #7144.8-16) for 2 days at room temperature. The bones were then washed for 2 days at room temperature using PBST (0.02% Triton-X100) and then incubated in secondary conjugated to Alexa-647 (Jackson Immuno Research, #711-606-152) for 2 days. Samples were then washed for 2 days and incubated in RIMS stepwise from 1.38-1.47 over 3 days. Bones were then imaged using a Zeiss 780 confocal.

### 3.10 Additional information

**Acknowledgements:** We thank Antti Lignell and Long Cai for technical help in building the LSFM and the Amgen Scientific Team (Frank Asuncion, David Hill, Mike Ominsky and Efrain Pacheco). Alon Greenbaum is a Good Ventures Fellow of the Life Sciences Research Foundation.

**Funding:** NIH Director's New Innovator IDP20D017782 and PECASE; Heritage Medical Foundation; Curci Foundation; Amgen-CBEA; Pew Charitable Trust; Kimmel Foundation; Caltech-COH.

**Author contributions:** H.J.McB. and V.G. conceived the project; A.G. and K.C. performed all experiments, data acquisition and analysis; T.D. and D.B. contributed computational tools and data analysis with input from R.B.; D.H.B. and H.M.K. contributed samples for clearing. A.G., K.C., T.D., D.B., and V.G. generated the figures and wrote the manuscript with input from all authors. V.G. supervised all aspects of the work.

**Competing interests:** V.G., K.C. and A.G. are inventors on patent application CIT-7683-P filled by California Institute of Technology that covers methods and devices for soft and osseous tissue clearing and fluorescent imaging. H.J.McB and R.B. are employees and shareholders of Amgen, Inc.

**Data and materials availability:** The r13c7 antibody is available from Amgen pending research proposal and material transfer agreement.

## Chapter 4

### **Engineered adeno-associated viruses for efficient and noninvasive gene delivery throughout the central and peripheral nervous systems**

[1] Chan, K.Y.C et al. “Engineered adeno-associated viruses for efficient and noninvasive gene delivery throughout the central and peripheral nervous systems”. In press.

#### **4.1 Summary**

Recombinant adeno-associated viruses (AAVs) are commonly used for *in vivo* gene transfer. Nevertheless, AAVs that provide efficient transduction across specific organs or cell populations are needed. Here, we used the capsid selection method CREATE to identify AAV-PHP.eB and AAV-PHP.S, two capsids that efficiently transduce the central and peripheral nervous systems, respectively. In the adult mouse, intravenous administration of  $1 \times 10^{11}$  vector genomes (vg) of AAV-PHP.eB transduced 69% of cortical and 55% of striatal neurons, while administration of  $1 \times 10^{12}$  vg AAV-PHP.S transduced 82% of dorsal root ganglion neurons, as well as cardiac and enteric neurons. The efficiency of these vectors facilitates stochastic, multicolor labeling of neurons, which can be used to assess individual cell morphology. To support such efforts, we provide methods for multicolor labeling of a tunable fraction of cells without compromising color diversity. Furthermore, we demonstrate that cell type-specific promoters can be used to expand the utility of these capsids across the nervous system. Collectively, these novel AAV tools enable efficient and versatile gene manipulation in the nervous system of transgenic and non-transgenic animals.

## 4.2 Introduction

Adeno-associated viruses (AAVs)<sup>14</sup> have been extensively used as vehicles for gene transfer to the nervous system enabling gene expression and knockdown, gene editing<sup>64,65</sup>, circuit modulation<sup>1,66</sup>, *in vivo* imaging<sup>67,68</sup>, disease model development<sup>69</sup>, and the evaluation of therapeutic candidates for the treatment of neurological diseases<sup>70</sup>. AAVs are well suited for these applications because they provide safe, long-term expression in the nervous system<sup>71,72</sup>. Most of these applications rely on local AAV injections into the adult brain to bypass the blood-brain barrier (BBB) and to temporally and spatially restrict transgene expression.

Targeted AAV injections have also been used for gene delivery to peripheral neurons to test strategies for treating chronic pain<sup>73-75</sup> and for tracing, monitoring, and modulating specific subpopulations of vagal neurons<sup>76,77</sup>. Many peripheral neuron populations, however, are difficult to access surgically (e.g., dorsal root ganglia (DRG), nodose ganglia, sympathetic chain ganglia, and cardiac ganglia) or are widely distributed (e.g., the enteric nervous system), thereby limiting methods for genetic manipulation of these targets. Likewise, in the CNS, single localized injections may be insufficient to study circuits in larger species<sup>78</sup> or to test gene therapies for diseases that involve the entire nervous system or widely distributed cell populations (e.g., Parkinson's, Huntington's, amyotrophic lateral sclerosis, Alzheimer's, spinal muscular atrophy, Friedreich's ataxia, and numerous lysosomal storage diseases)<sup>70</sup>.

Systemic AAV delivery<sup>79,80</sup> provides a non-invasive alternative for broad gene delivery to the nervous system; however, the high viral load required and relatively low transduction efficiency have limited wide adoption of this method. Several groups have developed AAVs that enhance gene transfer to the CNS after intravenous delivery. The recently reported AAV-AS capsid<sup>17</sup>, which utilizes a polyalanine N-terminal extension to the AAV9.47<sup>81</sup> VP2 capsid protein, provides higher neuronal transduction, particularly in the striatum, which may have applications for Huntington's disease. Similarly, the AAV-BR1 capsid<sup>16</sup>, based on AAV2, may be useful for applications that require more efficient and selective transduction of brain endothelial cells. Using a cell type-specific capsid selection

method we developed called CREATE (Cre REcombinase-based AAV Targeted Evolution), we recently reported AAV-PHP.B, a capsid that transduces the majority of neurons and astrocytes across many regions of the adult mouse brain and spinal cord after intravenous injection<sup>18</sup>. While the efficiency of AAV-PHP.B opens up new possibilities for CNS-wide genetic modification, it requires a substantial dose of vector (e.g.,  $1 \times 10^{12}$  vg per adult mouse or higher).

Here, we used CREATE to further evolve AAV-PHP.B for more efficient transduction of neurons throughout the adult mouse brain and spinal cord. We describe a novel enhanced variant of AAV-PHP.B, AAV-PHP.eB, which lowers the viral load required to transduce the majority of CNS neurons. Notably, we also report the characterization of a second capsid variant, AAV-PHP.S, that displays improved tropism towards peripheral neurons, including those in the DRG, cardiac ganglia, and enteric nervous system.

AAVs are also used for the bulk study of neuronal anatomical connectivity and morphology<sup>82</sup> and are components of multi-viral strategies for tracing the relationships between bulk inputs and outputs<sup>83,84</sup>. At the single-cell level, AAV-based multicolor labeling systems<sup>20</sup> have been developed with the goal of improving tracing efforts. However, the lack of control over the labeling density and uniformity of color diversity has been a persistent challenge<sup>20</sup>. To overcome these challenges, we have developed a two-component viral vector system to stochastically label cells with a wide range of hues while independently controlling the fraction of cells labeled. In addition, using the novel capsids reported here, we expressed a variety of fluorescent reporters under different cell type-specific promoters, supporting the potential use of these vectors for population-wide genetic manipulations of the nervous system in Cre transgenic or wild-type mice.

### 4.3 Results

#### Engineered AAV capsids for efficient transduction across the central and peripheral nervous systems

Here we report two novel AAV capsids that provide increased gene transfer to the CNS and the PNS via the vasculature. We developed these vectors by applying the previously published *in vivo* capsid selection method, CREATE<sup>18</sup> (**Fig. 1a**) to AAV-PHP.B and AAV9 as parental capsids. First, we sought to evolve the previously described AAV-PHP.B vector for more efficient tropism for the CNS. We generated an AAV capsid library in which the AAV-PHP.B 7-mer targeting sequence was modified by randomizing three consecutive amino acids (AAs) in an overlapping fashion across the 7-mer and flanking AAs (**Methods**). To provide selective pressure for capsids that transduce both neurons and glia, we subjected the library to parallel *in vivo* selections in adult Vglut2-IRES-Cre, Vgat-IRES-Cre, and GFAP-Cre mice, which express Cre in glutamatergic neurons, GABAergic neurons and astrocytes, respectively. After two rounds of selection, we found that a variant comprising the sequence *DGT* in place of the AAV-PHP.B sequence *AQTLAVPFK* at AA positions 587-589 (**Fig. 1b**) was enriched in all three transgenic lines (11.1%, 12.1%, and 15.7% of the total recovered sequences, respectively). We used this DGT substitution variant or AAV-PHP.B to package a single-stranded (ss) AAV reporter genome that expresses a nuclear-localized GFP (NLS-GFP) from the constitutive CAG promoter, (ssAAV-CAG-NLS-GFP), and delivered  $1 \times 10^{11}$  vg of the viruses by intravenous injection to adult mice. We found that the DGT substitution variant appeared to transduce the CNS more efficiently than AAV-PHP.B as judged by wide-field fluorescence microscopy of the intact brain (**Fig. 1c**) and confocal microscopy on thin sections from brain (**Fig. 1f**) and spinal cord (**Fig. 1g**). We refer to this enhanced AAV-PHP.B variant as AAV-PHP.eB and present further quantitative characterization below. Second, we performed a separate selection of an AAV9-based 7-mer library in GFAP-Cre mice<sup>18</sup>. After two rounds of selection, we found a variant (7-mer sequence – QAVRTSL, representing 33.3% of the total sequences recovered; **Fig. 1b**). As above, we evaluated the capsid against AAV9, its parent capsid, by packaging the ssAAV-

CAG-NLS-GFP reporter and delivering the vectors at  $1 \times 10^{12}$  vg/mouse. While the QAVRTSL variant and AAV9 both showed similar sparse transduction of the brain (**Fig. 1c, d**), the QAVRTSL variant notably appeared to strongly transduce peripheral sensory afferents entering the spinal cord and brain stem (**Fig. 1c-e**). We refer to this variant as AAV-PHP.S. While unexpected given the selection scheme, the improved tropism for sensory neurons provided by AAV-PHP.S could be useful for widespread gene transfer to cells of the PNS. We explore this possibility below.

Together these data indicate that the tropism of AAV9 can be routed towards either more efficient CNS or PNS transduction (**Fig. 1c**), depending on the AA insertion/modification. A comparison of the AA insertion/modification sequences of AAV-PHP.eB and AAV-PHP.S along with their parent capsids, AAV-PHP.B and AAV9, respectively, is shown in **Fig. 1b**, and their nucleotide sequences and relative viral production efficiencies are provided in (**Supplementary Table 1**).

### **Quantification of transduction by AAV-PHP.eB and AAV-PHP.S**

Next, we quantified the transduction efficiency of AAV-PHP.eB and AAV-PHP.B three weeks after intravenous administration of  $1 \times 10^{11}$  vg of ssAAV-CAG-NLS-GFP. First, when compared with AAV-PHP.B, AAV-PHP.eB provided significantly increased ( $> 2.5$ -fold) GFP fluorescence per cell nucleus throughout the cortex and striatum and ( $> 1.8$ -fold) in cerebellar Purkinje cells (**Fig. 2d** and **Supplementary Fig. 1**). To quantify expression in Purkinje cells, an automated cell segmentation and analysis pipeline was developed (**Supplementary Fig. 2** and **Methods**). AAV-PHP.eB also showed an increase over AAV-PHP.B in both the percentage of total cortical cells transduced (DAPI<sup>+</sup> cells that express GFP;  $51 \pm 2\%$  vs.  $37 \pm 3\%$ ; **Fig. 2a,e**) and the percentage of NeuN<sup>+</sup> neurons transduced ( $69 \pm 4\%$  vs.  $49 \pm 7\%$ ; **Fig. 2b,e**). In the striatum, a similar trend towards increased transduction by AAV-PHP.eB compared to AAV-PHP.B was observed from quantifying both the percentage of DAPI<sup>+</sup> and NeuN<sup>+</sup> cells that express GFP (DAPI<sup>+</sup>:  $42 \pm 5\%$  vs.  $26 \pm 5\%$ ; **Fig. 2a,e**; NeuN<sup>+</sup>:  $55 \pm 7\%$  vs  $36 \pm 8\%$ , **Fig. 2b,e**). In the cerebellum, we observed an increase in

the percentage of Calbindin<sup>+</sup> Purkinje cells that expressed GFP in mice injected with AAV-PHP.eB as compared with AAV-PHP.B ( $76 \pm 5\%$  vs.  $37 \pm 10\%$ , **Fig. 2c,e**). Both vectors transduced a similar percentage of S100<sup>+</sup> glia in the cortex as well as the striatum (**Fig. 2e** and **Supplementary Fig. 3**).

Next, we examined the transduction efficiency of AAV-PHP.S compared with AAV9, a vector known to provide broad transduction of the PNS and many peripheral organs after intravenous delivery<sup>79,85-87</sup>. We packaged ssAAV-CAG-NLS-GFP into AAV9 or AAV-PHP.S, administered the viruses to adult mice by intravenous injection at  $1 \times 10^{12}$  vg/mouse, and then assessed GFP expression after three weeks. In the DRG, AAV-PHP.S provided a significant increase in transduction over AAV9 (**Fig. 3a**) as quantified both by native GFP fluorescence per cell nucleus ( $> 2.6$ -fold) (**Fig. 3b**) and by the percentage of cells positive for PGP9.5, a marker for neurons in the PNS, that co-expressed GFP ( $82 \pm 2\%$  vs.  $46 \pm 0.7\%$ ) (**Fig. 3c**). AAV-PHP.S also transduced the cardiac ganglia (**Fig. 3a**) and cells within the enteric nervous system of the small intestine and colon (**Fig. 3d** and **Supplementary Movie 1**) with high efficiency. To determine the identity of enteric nervous system cells transduced by AAV-PHP.S, we performed immunohistochemistry (IHC) for the neuronal marker PGP9.5 and the astrocyte marker S100b. The majority of GFP<sup>+</sup> cells within the myenteric plexus (**Fig. 3d**) and submucosal plexus (**Supplementary Fig. 4**) of the small intestine and colon co-localized to PGP9.5<sup>+</sup> cells with sparse and faint transduction seen in S100b<sup>+</sup> cells (**Supplementary Movie 1**). Notably, we also observed robust transduction by AAV-PHP.S of cells in many organs, including the liver, lungs, heart, and stomach cells within the muscle and myenteric plexus layers (**Supplementary Fig. 4**).

### **AAV-PHP.eB and AAV-PHP.S enable efficient co-transduction of neurons**

The increased transduction efficiency and enhanced expression per cell observed with both AAV-PHP.eB and AAV-PHP.S, relative to their parental capsids, suggested that a large fraction of cells are transduced by multiple AAV genomes when administered at the above doses. If true, the transduction efficiency of these vectors may enable applications that require expression from more than one AAV genome. To test this, we co-administered a cocktail of three AAV genomes with the single fluorescent proteins (XFPs) mRuby2,

mNeonGreen, and mTurquoise2 expressed from the human Synapsin (hSyn1) promoter, which limits expression to neurons. The three-vector mix (AAV-hSyn1-3XFP) was packaged into AAV-PHP.eB or AAV-PHP.S and injected into adult wild-type mice. Co-administration of these vectors resulted in dense labeling of cells with multiple colors in the brain when delivered with AAV-PHP.eB and in the enteric nervous system with AAV-PHP.S (**Fig. 4a**). Because cells transduced by a single vector would express only one of the three XFPs, and thereby be red, green or blue, the wide color diversity seen in both the brain and the enteric nervous system demonstrates that these novel capsids enable the independent and stochastic co-transduction of neurons by multiple viral vectors.

### **Tunable density multiplexed gene expression**

Multicolor labeling by AAVs can benefit single-cell morphology and tracing studies<sup>88</sup>. However, direct local virus injections<sup>20,82</sup> may not provide uniform color diversity because of the non-uniform distribution of vector copies radiating outward from the injection site<sup>20</sup>. We predicted that systemically delivered AAVs may offer a solution to this challenge because the distribution of labeled cells and color diversity would be more uniform. Indeed, when we delivered three separate XFPs expressed from the CAG promoter (AAV-CAG-3XFP; **Fig. 4b**) we observed a wide range of hues in the cortex, striatum, and cerebellum (**Fig. 4d, left**). However, this strategy, where the expression of the transgene is controlled *in cis*, couples the expression level, labeling density, and the color diversity to the virus dose. As expected, when we lowered the virus dose we observed that the density and color diversity were reduced (i.e., most cells were labeled by primary colors – red, green, or blue – likely reflecting single transduction events) (**Fig. 4d, right**). Therefore, to achieve stochastic, multicolor labeling while reducing the fraction of labeled cells, we sought a two-component AAV system in which the expression of XFPs in each cell is dependent on a separate inducer vector that is co-transduced at a variable dose. In this way, color diversity could be decoupled from the labeling density. We developed two inducer systems: one that uses the tetracycline (tet)-inducible system – a tet-off transactivator (tTA) and a tet-responsive element (TRE)<sup>89</sup> (**Fig. 4c**) – and a second that uses the FLPo-FRT system<sup>90</sup> (**Supplementary Fig. 5**).

Importantly, both systems should allow for cell type-specific expression through the use of FLEX/DIO vectors and the extensive collection of Cre transgenic animals available (e.g., JAX, ABI, and GENSAT).

The tTA-TRE based system relies on co-administration of three separate vectors with XFP expression driven by a TRE containing promoter, along with an adjustable dose of an AAV-tTA expression vector. To evaluate this approach, we co-administered ssAAV-TRE-mNeonGreen along with a ssAAV-CAG-tTA vector. We delivered ssAAV-TRE-mNeonGreen at the same dose across all animals while varying the dose of the ssAAV-CAG-tTA inducer. We observed a dose-dependent reduction in the density of cells expressing mNeonGreen two weeks after intravenous delivery of the two vectors (**Supplementary Fig. 6**). Similar results were achieved with the two-component FLPo-FRT system (**Supplementary Fig. 5**). Although both systems are viable, we focused on the tTA-TRE system because (1) the short (417 bp) TRE promoter is compatible with longer transgenes and (2) the added versatility of providing higher levels of expression with a positive-feedback loop in the AAV-tTA vector (**Supplementary Fig. 7**). However, care should be taken with long-term use of the tTA-TRE overexpression system to avoid the potential toxicity seen with overexpression of the tTA<sup>91</sup>.

We next sought to test the tTA-TRE system with multi-color labeling. Using the same 3XFP cocktail described above (**Fig. 4a**), we generated three AAV-TRE-XFP genomes and delivered the vectors at a high dose to all animals, along with either  $1 \times 10^{11}$  or  $1 \times 10^{10}$  vg/animal of a neuron specific inducer vector (ssAAV-PHP.eB:hSyn1-tTA). Three weeks after intravenous administration we observed robust and inducer dose-dependent labeling (**Fig. 4e**). Even at the lower inducer dose, and in contrast with the single-component expression system (**Fig. 4d**), many cells expressed more than one of the three XFP transgenes. We quantified the fraction of cells that contained each XFP and found that, with the two-component system at a low inducer dose, at least 42% of the labeled cells within the cortex, striatum, and the Purkinje cell population expressed all three XFPs. In comparison, in the one-component system, less than 10% of the labeled cells in the cortex and 1% of cells in the striatum and the Purkinje cell population expressed all three XFPs (**Fig. 4f**). Finally,

we utilized the two-component system to trace mitral cells in the olfactory bulb of an adult Tbx21-Cre mouse, a transgenic line that labels mitral and tufted cells in the olfactory bulb. Three weeks after intravenous administration of  $1 \times 10^{12}$  vg/animal of a cocktail of TRE-DIO-XFPs and  $1 \times 10^{10}$  vg/animal of AAV-ihSyn1-tTA, a positive feedback inducible tTA driven by the hSyn promoter, we observed multicolor labeling at a density that facilitated segmentation of several local dendritic arbors after tissue clearing (**Fig. 4g**).

### **A versatile AAV toolbox with cell-type specific promoters**

A further advantage of systemic AAV administration is that it provides a means to transduce populations of specific cell types even if the cells are distributed broadly. As a proof-of-concept that AAV-PHP.eB can be used to achieve cell type-restricted expression across the brain, we co-administered three ssAAV genomes that each expressed a spectrally distinct XFP controlled by a different cell type-specific promoter. In this way, we were able to independently express distinct transgenes in neurons (hSyn1), astrocytes (GFAP), and oligodendrocytes (MBP) within the same animal (**Fig. 5a**). We also packaged AAV vectors containing previously published promoters that restrict reporter expression to catecholaminergic cells (mouse tyrosine hydroxylase; mTH), serotonergic cells (FEV promoter; Ple67), Purkinje cells (PCP2 promoter; Ple155)<sup>92</sup>, or GABAergic interneurons within the forebrain (mouse distal-less homeobox; mDlx5/6)<sup>93</sup> and administered these intravenously (**Fig. 5b-f**). We provide quantification of the specificity and efficiency of reporter transduction using these gene regulatory elements in **Figure 5f**.

Similarly, to achieve cell type-restricted expression within the PNS, we used AAV-PHP.S to co-administer the 3XFP cocktail under the control of the hSyn1 promoter (ssAAV-PHP.S:hSyn-3XFP) or by expressing the 3XFP cocktail from the CAG promoter in a double-floxed inverted open reading frame (ssAAV-PHP.S:CAG-DIO-3XFP) for expression in cholinergic neurons when used with ChAT-IRES-Cre transgenic mice (**Supplementary Fig. 8**). Using this strategy together with the two-component TRE-tTA system described above, we were able to sparsely label cholinergic neurons within the myenteric plexus in the

proximal colon and trace the processes of multiple neurons for several millimeters (**Supplementary Fig. 8**).

In **Supplementary Table 2**, we summarize the available constructs presented here, along with their applications within this paper.

#### 4.4 Discussion

Using the previously described CREATE<sup>18</sup> selection method, we have developed and characterized two novel capsids, AAV-PHP.eB and AAV-PHP.S, that enable efficient and noninvasive gene delivery throughout the CNS or PNS and visceral organs, respectively (**Fig.1 and Supplementary Fig. 4**). As in our prior study<sup>18</sup>, these variants were identified after only two rounds of *in vivo* selection, further demonstrating how CREATE can be used to identify AAV capsids with enhanced transduction properties.

Previously, we selected for AAV capsids that more efficiently crossed the BBB and transduced GFAP-expressing astrocytes. Several of the identified variants, most notably AAV-PHP.B, also transduced neurons, although in general not with the same efficiency. In the current work, we sought to improve neuronal transduction while maintaining strong astrocytic transduction by performing parallel selections for capsids that transduced two widely distributed glutamatergic and GABAergic neuronal populations as well as astrocytes. This led to the development of the AAV-PHP.eB capsid, a variant that was consistently enriched through selections in each of the three Cre lines. Remarkably, AAV-PHP.eB differs from AAV-PHP.B at only two AAs adjacent to the original 7-mer insertion, yet at the relatively low systemic dose of  $1 \times 10^{11}$  vg/mouse it transduces the majority of neurons in the cortex and striatum, as well as over 75% of cerebellar Purkinje cells. The success of this approach suggests that the continued evolution of the AAV-PHP.B-based capsid using CREATE in multiple Cre lines is a viable strategy to improve transduction of other cell types, such as oligodendrocytes, while maintaining the overall CNS tropism.

We also describe the AAV-PHP.S capsid, which we identified by selecting for capsids that transduce GFAP-expressing cells. Although the mechanism underlying the

selection of this variant is not clear, its discovery is nevertheless important; to date, there have been no reports of recombinant vectors capable of efficient, widespread transduction of peripheral neurons via systemic delivery in adult animals. Consequently, studies of PNS circuitry are arguably lagging behind those in the CNS. Many peripheral neuronal populations are challenging to access surgically, especially in small rodents, owing to their location<sup>94</sup> or broad distribution<sup>95</sup>. For example, the DRG are a key target for gene therapy or genetic manipulation in the study of pain<sup>74,75</sup>. Our finding that AAV-PHP.S transduces at least 80% of DRG neurons should enable high coverage delivery of transgenes to multiple DRG. More broadly, AAV-PHP.S also transduces neurons in the enteric nervous system and in other peripheral ganglia, such as the cardiac ganglia, and could provide system-wide gene delivery to these populations, or to subsets of these populations when used with Cre transgenic lines or cell type-specific promoters. Encouragingly, several recent studies have demonstrated that the tools developed for studying the brain can also be applied to probe the function and projections of specific cell types within the PNS<sup>73,76,77</sup>. The importance of anatomically and functionally mapping the PNS has been recognized by the recent NIH initiative for Stimulating Peripheral Activity to Relieve Conditions (SPARC), and the novel tropism of AAV-PHP.S may bring about new strategies to modulate PNS circuits through genetic tools such as optogenetics<sup>1</sup> or chemogenetics<sup>66</sup>. For example, systemic Flp- and Cre-dependent effectors delivered via AAV-PHP.S to Cre transgenic animals or used in combination with circuit-specific retrograde Cre viruses should allow for circuit-based interrogation of PNS function and enable the types of experiments now routinely conducted in the CNS<sup>1,96</sup>.

The improved transduction efficiency we observed from both AAV-PHP.eB and AAV-PHP.S relative to their parental capsids could be seen from both the increased number of transduced cells and the increased expression level per cell (**Fig 2, 3** and **Supplementary Fig. 1**). The higher expression in individual cells likely results from an increased mean number of vector genomes per cell, as supported by the high frequency of cells that were labeled with multiple XFPs (**Fig. 4a**). This suggests that for applications that do not require high expression levels per cell, the viral doses could be lowered from those used in this study.

Lowering the viral load used for systemic gene delivery will reduce cost and production burden<sup>14</sup> and minimize the risk for adverse reactions to the viral components<sup>97,98</sup>.

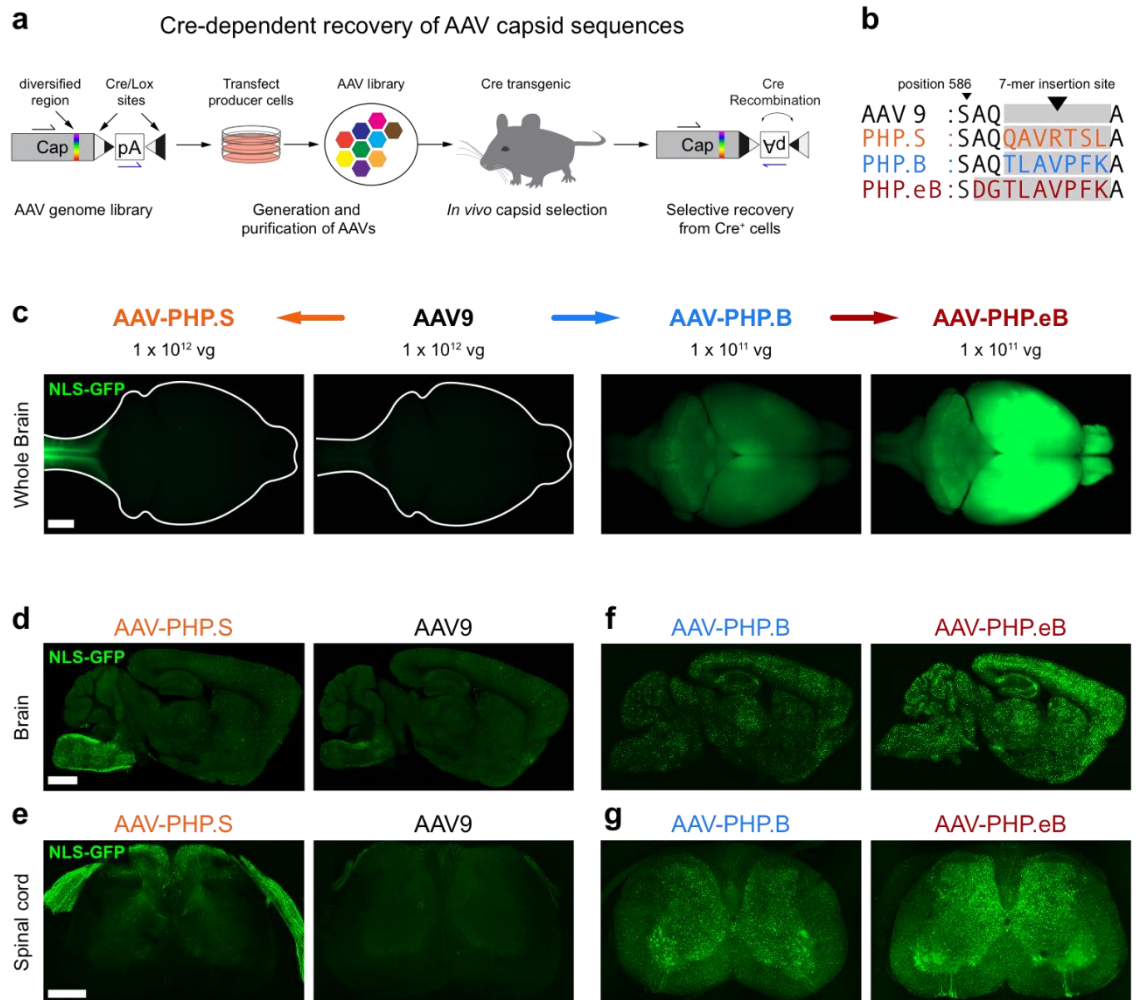
As AAV vector engineering efforts continue to improve the efficiency of transduction of specific organs or cell populations<sup>99</sup>, widespread gene transfer will become a more viable alternative to the development of germline transgenics. To facilitate studies in nontransgenic animals, it will be necessary to develop additional vehicles with improved transduction efficiency and specificity and to identify gene regulatory elements that can be used to restrict expression to defined populations. Finding short enhancers that provide the expected restriction in expression has proven challenging<sup>100</sup>, but more recent efforts<sup>92</sup> are expanding the collection of enhancer-promoters available. Here, we evaluated several gene regulatory elements and found that it was possible to achieve expression that was largely restricted to the expected cell types. If the separate vector components are translatable across species, vector-based genetic manipulations of specific cell types may become more feasible in genetically intractable organisms. Notably, in a recent study, intravenously administered AAV-PHP.B provided improved transduction of the rat CNS, relative to AAV9<sup>101</sup>. In a recent study, a *Dlx5/6* enhancer was found to drive expression that was largely restricted to GABAergic neurons in the telencephalon of a panel of vertebrate species<sup>93</sup>. In this study, we further validate the *Dlx5/6* enhancer vector for forebrain-wide GABAergic neuron expression when delivered using intravenous AAV-PHP.eB (**Fig. 5e**).

Cell morphology both informs and determines cell function<sup>102,103</sup>, and numerous tools have been developed for morphology and connectivity applications<sup>104,105</sup>, especially in the brain<sup>106,107</sup>. To enable tracing of cellular processes, sparse labeling of cells has been a key enabler – whether this comes from restricted gene or dye delivery (e.g. single-cell electroporation<sup>108,109</sup>) or from stochastic multicolor labeling and color-channel separation<sup>20,106,107</sup>. Both viral and transgenic approaches have been employed to address this problem. Although transgenic mice with sparse gene expression<sup>110,111</sup> are useful, AAVs<sup>20</sup> provide a versatile platform that can be quickly customized with up-to-date tools, promoters, and recombinase-dependent expression control systems. Furthermore, vectors can be delivered through different routes or at a specific dose to control the location and fraction of

cells labeled. For these reasons, we developed an AAV-based expression system that allows for sparse multicolor labeling. This system uses two AAV components: a cocktail of XFP encoding vectors that exist in the off-state until induced by a second AAV vector that can be delivered at an empirically determined dose to provide expression in the desired fraction of co-transduced cells. We demonstrated that this system is well suited to the use of systemically delivered AAV-PHP.eB and AAV-PHP.S to achieve relatively uniform and sparse expression in the CNS and PNS. In addition, decoupling the relationship between the labeling density and the expression level may have additional applications. For example, sparse, high-level expression could be beneficial for *in vivo* monitoring of neural activity with genetically encoded activity indicators by lowering the background fluorescence during wide-field imaging<sup>112</sup> or large area 2-photon imaging<sup>113</sup>. Furthermore, a similar framework could be applied to the understanding of how neuronal circuitry and morphology change during normal and diseased states. Future applications of this two-component approach could also be used to achieve intersectional gene expression in defined populations through circuit-specific delivery of one or both components or through the use of unique capsid-promoter or capsid-capsid combinations that serve to restrict expression to more specific populations.

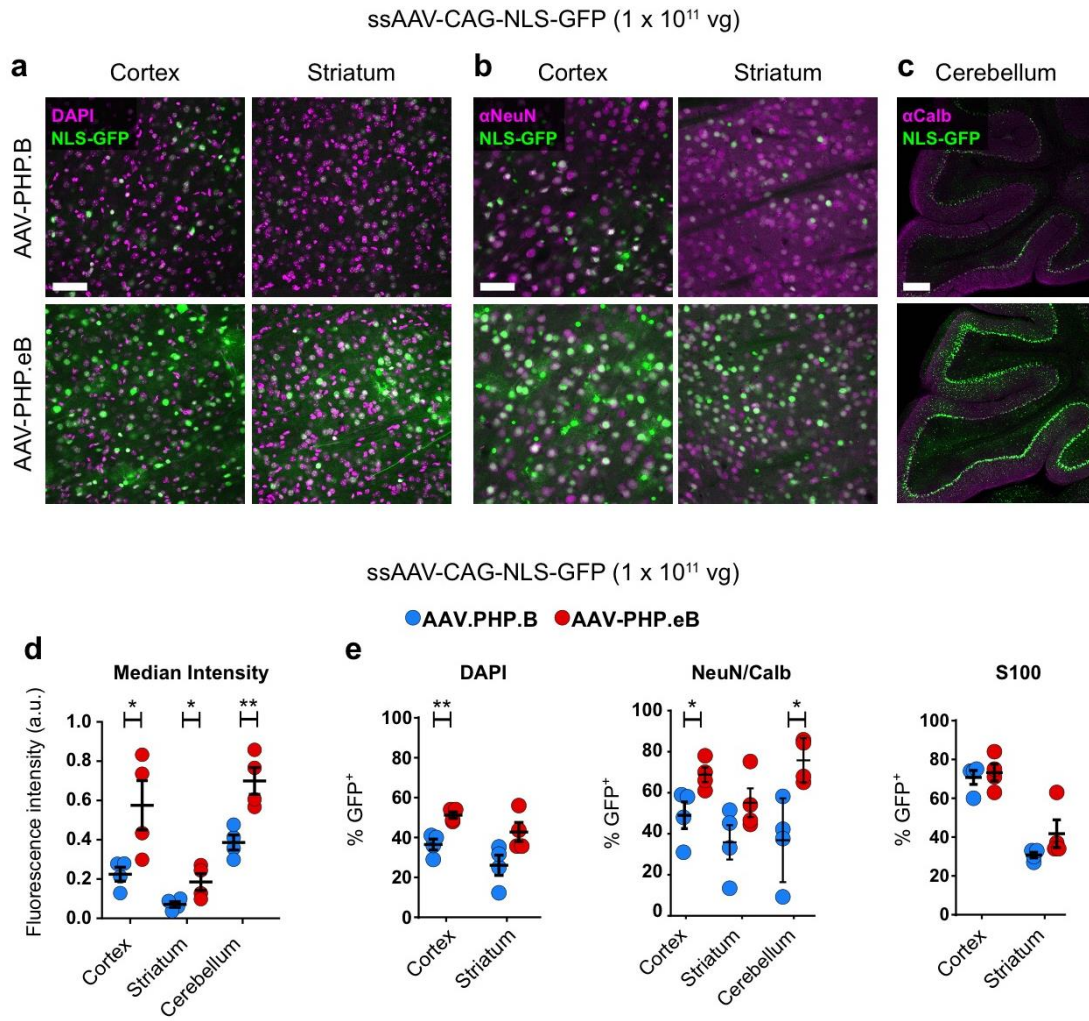
Overall, the vector resources we present here expand the AAV toolbox for the genetic modification of neurons with the central and peripheral nervous systems.

## 4.5 Main Figures



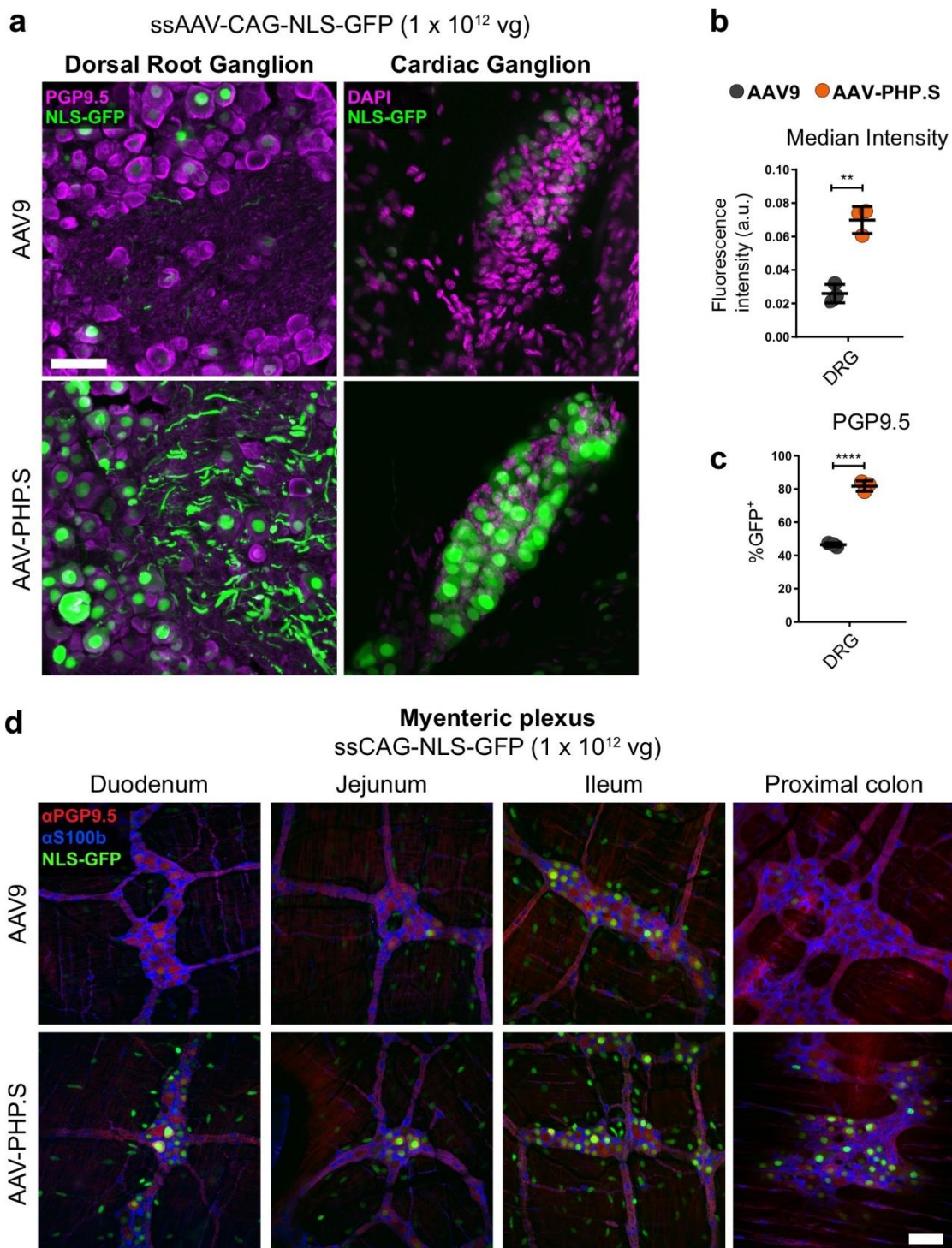
**Figure 1. Engineered AAV capsids for efficient transduction across the peripheral and central nervous systems.** (a) Schematic of the CREATE selection method. (b) The amino acid (AA) sequences for the 7-mer insertions and flanking sequences for AAV-PHP.S, .B and .eB; the 7-mer and adjacent substitutions are highlighted in colored text. (c-g) ssAAV-CAG-NLS-GFP was packaged into the indicated capsid and intravenously injected into adult mice at  $1 \times 10^{12}$  vg/mouse (AAV9 and AAV-PHP.S; c-e) or  $1 \times 10^{11}$  vg/mouse (AAV-PHP.B and AAV-PHP.eB; c, f, and g). (c) Representative whole brains imaged by wide-field fluorescence microscopy, with similar imaging and display parameters across all four brains.

Imaging was performed after three weeks of expression. **(d-g)** Representative confocal images of native GFP fluorescence from sagittal brain sections (**d** and **f**) and transverse spinal cord sections (**e** and **g**) are shown for the indicated capsids. For **(d-g)**, all imaging and display conditions are matched across panel pairs to allow for comparisons. Panels **(d)** and **(f)** are 40  $\mu\text{m}$  maximum intensity projections (MIPs) and panels **(e)** and **(g)** are 300  $\mu\text{m}$  MIPs. Scale bars for **(c-e)** are 1 mm.

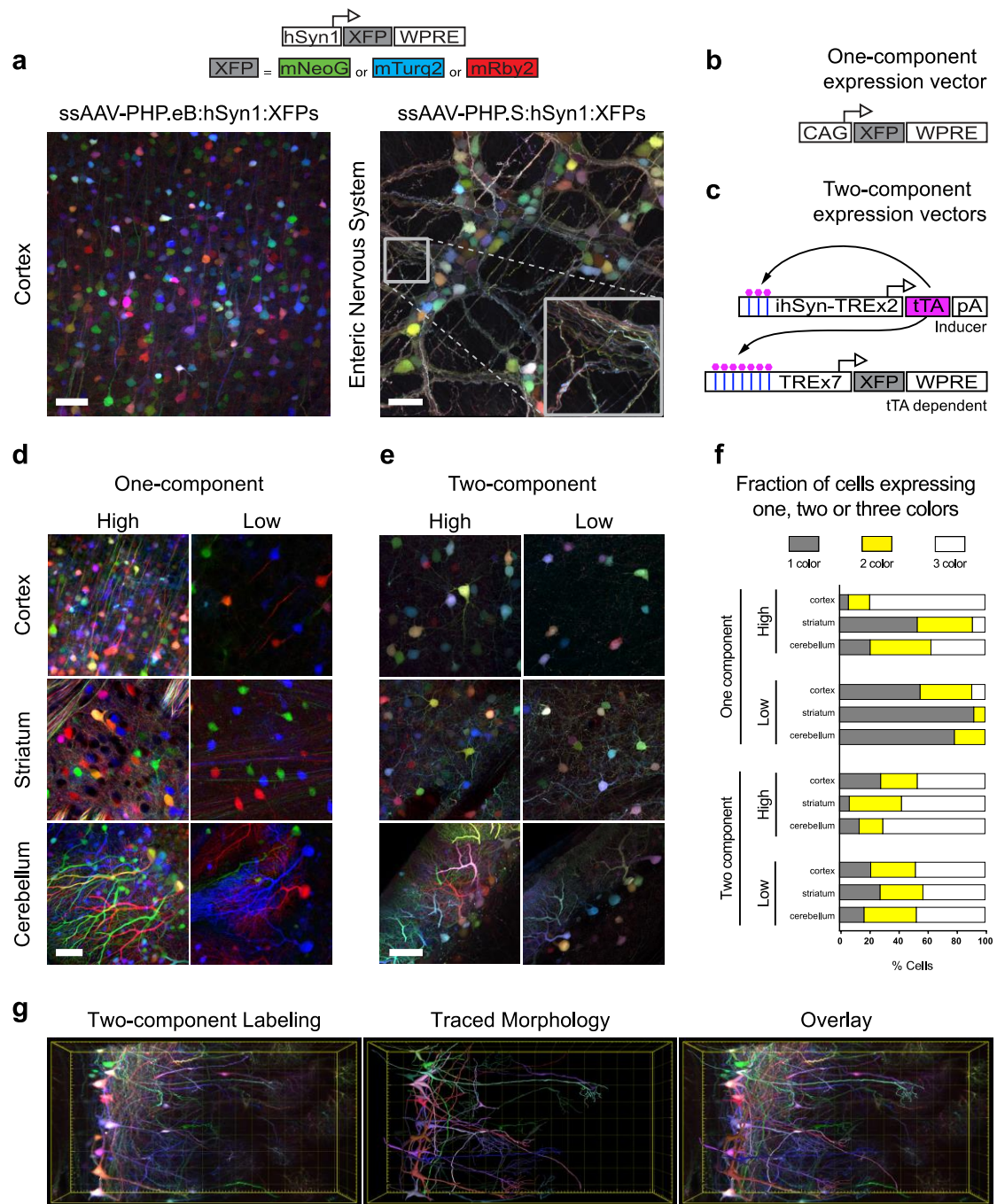


**Figure 2. AAV-PHP.eB transduces several CNS regions more efficiently than AAV-PHP.B.** ssAAV-PHP.B:CAG-NLS-GFP or ssAAV-PHP.eB:CAG-NLS-GFP was intravenously injected into adult mice at  $1 \times 10^{11}$  vg/mouse. Native GFP fluorescence was assessed after three weeks of expression. (**a** and **b**) Representative images of native GFP fluorescence (green) and DAPI staining (**a**, magenta) or NeuN staining (**b**, magenta) in the cortex and striatum. (**c**) Representative images of Calbindin immunohistochemistry (IHC) (magenta) in the cerebellum. (**d**) The mean of the median GFP intensity in all DAPI<sup>+</sup> cells in the cortex and striatum or in Calbindin<sup>+</sup> cells in the cerebellum. (**e**) Quantification of cell transduction. From left to right: The percentage of DAPI<sup>+</sup> cells that express GFP in the

cortex:  $51 \pm 2\%$  and  $37 \pm 3\%$  for AAV-PHP.eB and AAV-PHP.B, respectively ( $p = 0.0034$ ) and striatum:  $42 \pm 5\%$  and  $26 \pm 5\%$  for AAV-PHP.eB and AAV-PHP. B, respectively ( $p = 0.0541$ ). The percentage of NeuN<sup>+</sup> cells that express GFP in the cortex:  $69 \pm 4\%$  and  $49 \pm 7\%$  for AAV-PHP.eB and AAV-PHP.B, respectively ( $p = 0.037$ ) and in the striatum:  $55 \pm 7\%$  and  $36 \pm 8\%$  for AAV-PHP.eB and AAV-PHP.B, respectively (n.s.,  $p = 0.128$ ). The percentage of Calbindin<sup>+</sup> cerebellar Purkinje cells that express GFP:  $76 \pm 5\%$  and  $37 \pm 10\%$  for AAV-PHP.eB and AAV-PHP. B, respectively ( $p = 0.039$ ). The percentage of S100<sup>+</sup> cells that express GFP in the cortex and striatum ( $p = 0.6738$  and  $0.1814$ ). **(d and e)** For quantification:  $n = 4$  mice per group, mean  $\pm$  SEM, unpaired, parametric t-test ( $***p \leq 0.001$ ;  $**p \leq 0.01$ ;  $*p \leq 0.05$ ; n.s.,  $p \geq 0.05$ ). All images were single-plane confocal images of native GFP fluorescence. Scale bars are  $50 \mu\text{m}$  **(a and b)** and  $100 \mu\text{m}$  **(c)**.

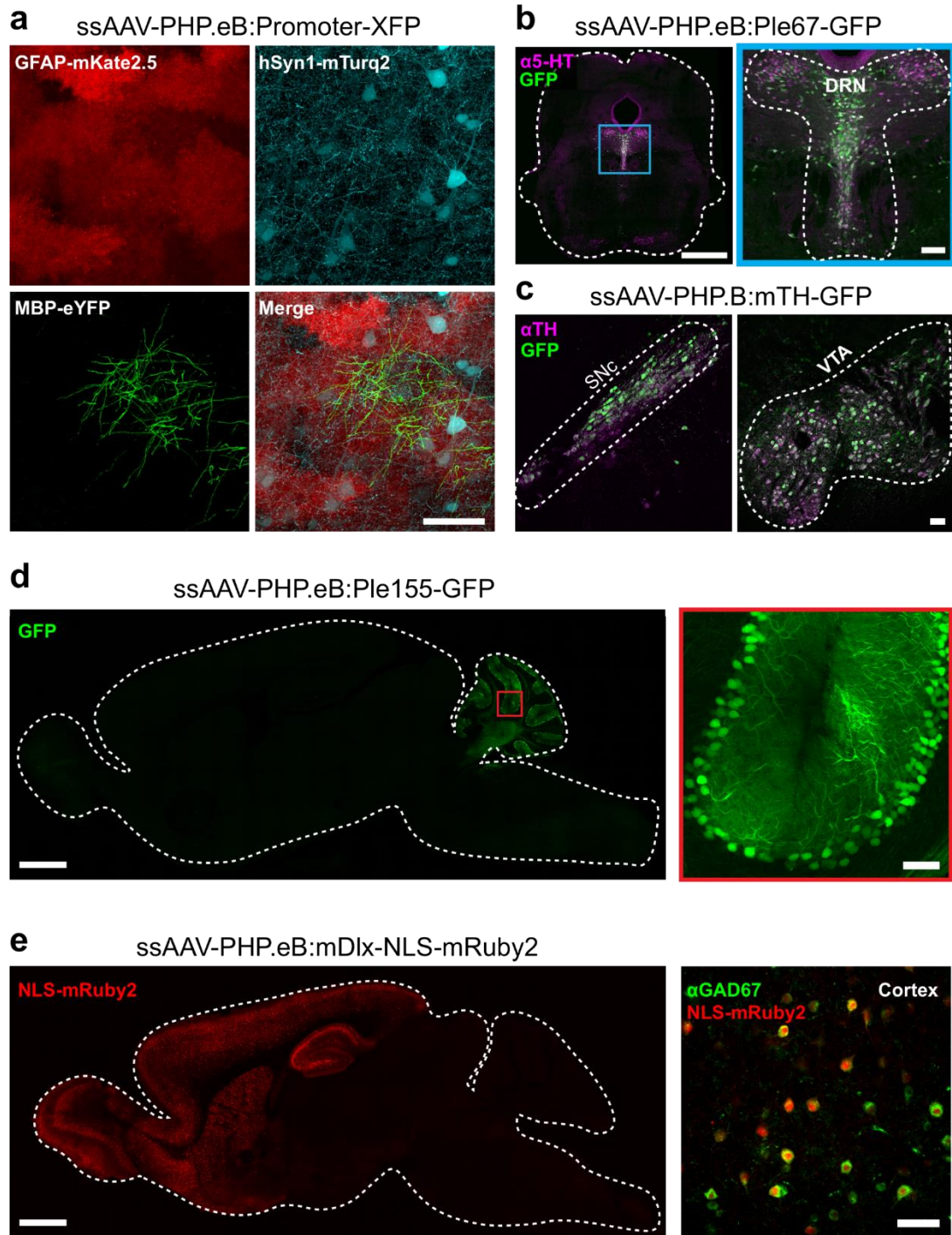


**Figure 3. AAV-PHP.S efficiently transduces peripheral neurons.** ssAAV-PHP.S:CAG-NLS-GFP or ssAAV9:CAG-NLS-GFP was intravenously injected into adult mice at  $1 \times 10^{12}$  vg/mouse. Native GFP fluorescence was assessed after three weeks of expression. **(a)** Representative images of GFP expression and neuronal PGP9.5 (left, magenta) or DAPI (right, magenta) staining in a dorsal root ganglion (DRG, left) and cardiac ganglion (right). **(b)** Quantification of the mean of the median GFP fluorescence intensity per cell with AAV-PHP.S or AAV9. **(c)** Quantification of the percentage of PGP9.5<sup>+</sup> cells transduced by AAV-PHP.S and AAV9; mean  $\pm$  SEM. **(d)** Representative images of GFP expression, with neuronal PGP9.5 (red) and astrocyte S100b (blue) staining in the myenteric plexus of the duodenum, jejunum, ileum and proximal colon. In **(b)** and **(c)** means were compared with unpaired parametric t-tests;  $n = 3$  independent animals per group. All imaging and display conditions are matched in the GFP channel across panel pairs. Scale bars in **(a)** and **(d)** are 50  $\mu$ m.



**Figure 4. AAV-PHP.eB and AAV-PHP.S transfer multiple genomes per cell and enable tunable multicolor labeling with a two-component system.** (a) Co-administration of a cocktail of three ssAAV vectors that each express one of three XFPs (top, schematic) from the hSyn1 promoter labels individual cells with a range of hues. Representative images are

shown of the cortex (left, AAV-PHP.eB; total dose:  $4 \times 10^{11}$  vg) and the enteric nervous system (right, AAV-PHP.S; total dose:  $1 \times 10^{12}$  vg). The inset box in gray is a high magnification view that highlights the dense network of discernable individual neurites labeled with a range of hues **(b)** Schematic of a one-component system where expression of the transgene is dependent on a *cis* regulatory element. **(c)** Schematic of the two-component system based on the tTA-TRE inducible system. Expression from the TRE-XFP vectors is dependent on co-transduction with a tTA (tet-off) inducer vector that contains a positive feedback loop to increase tTA expression (ssAAV-PHP.eB:ihSyn-tTA). The tTA protein (magenta hexagons) binds to TRE elements (blue lines) present on both the inducer and XFP vector. The dosage of the inducer vector can be varied, allowing sparse or dense gene expression. **(d)** Representative images of the cortex, striatum, and cerebellum using the one-component expression system with a 3XFP cocktail expressed from the CAG promoter at an equal dose per XFP (*high*: total dose =  $2 \times 10^{11}$  vg; *low*: total dose =  $1 \times 10^{10}$  vg). **(e)** Representative images of the cortex, striatum, and cerebellum using the two-component system. The co-administered cocktail of viruses included  $1 \times 10^{12}$  total of the TRE-3XFP cocktail and a high ( $1 \times 10^{11}$  vg) or low ( $1 \times 10^{10}$  vg) dose of the tTA inducer. **(f)** Percentages of individual cells in each condition that express one (gray), two (yellow), or all three (white) XFPs. **(g)** MIP (left) of a coronal section from the olfactory bulb of a Tbx21-Cre mouse transduced with a cocktail of ssAAV-PHP.eB:TRE-DIO-3XFP ( $1 \times 10^{12}$  vg total dose) and ssAAV-PHP.eB:ihSyn-tTA ( $1 \times 10^{10}$  vg). The morphologies of local neuronal dendritic arbors were traced with NeuTube and imported into Imaris v8.3 for 3D rendering (middle). An overlay is provided (right). Traced neurons are color-matched across panels in **(g)**. Scale bars **(d and e)** are 50  $\mu\text{m}$  and are MIPs of 40  $\mu\text{m}$  z-stacks. The total volume in **(g)** is  $0.042\text{mm}^3$ . All images were acquired on tissue extracted three weeks after intravenous delivery in adult mice, except for **(a)** where the expression was assessed after 11 days.

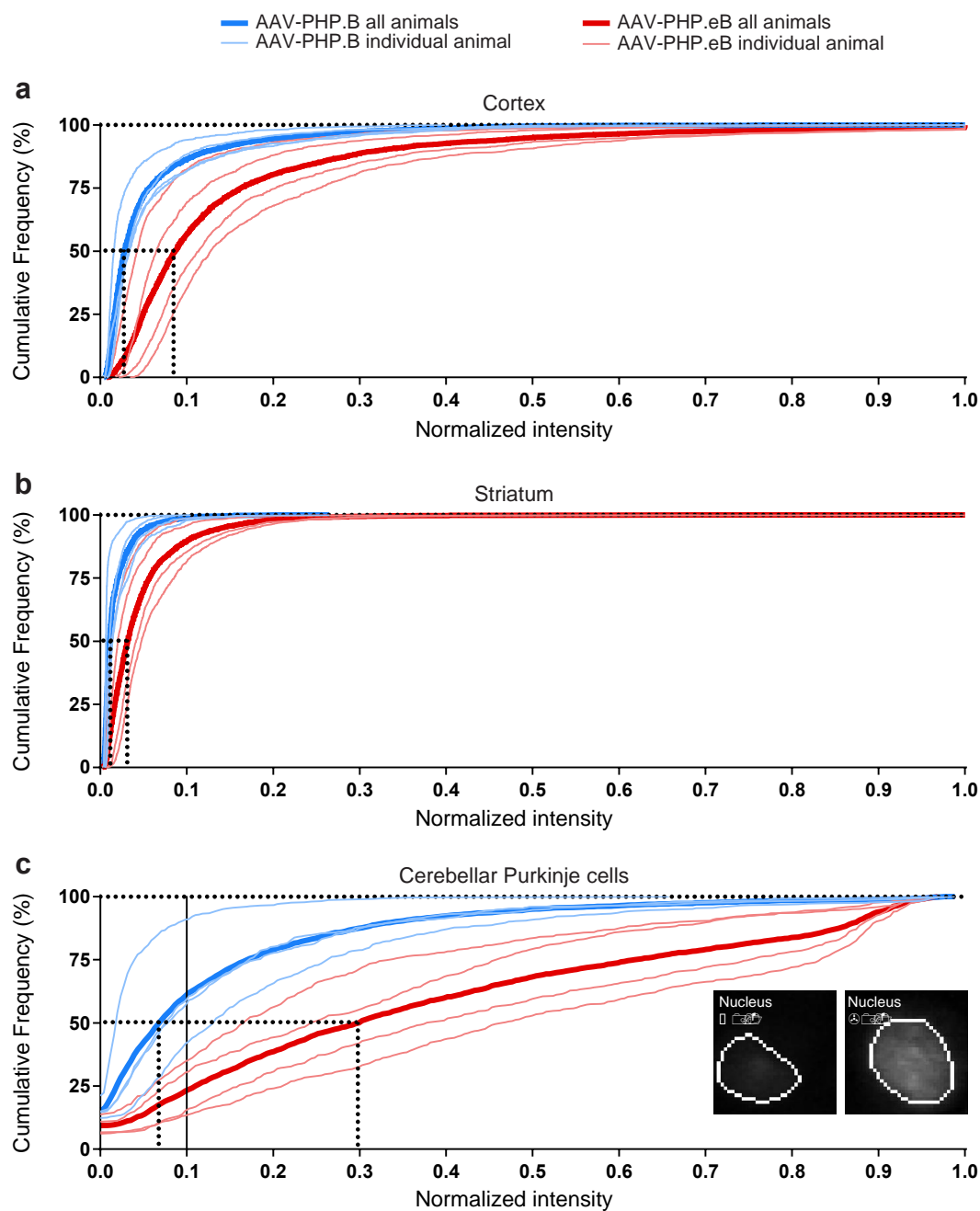


**f**

Promoter	Efficiency (%)			Specificity (%)			Length (bp)	Dosage (vg)
FEV/Plc67	DRN: 67	MRN: 48		DRN: 89	MRN: 55		2202	1 x 10 <sup>12</sup>
Mouse tyrosine hydroxylase	DRN: 57	SNc: 99	VTA: 86	DRN: 22	SNc: 81	VTA: 81	2571	1 x 10 <sup>12</sup>
mDlx w/ chimeric intron	CTX: 81	HIP: 90		CTX: 87	HIP: 82		530	3 x 10 <sup>11</sup>

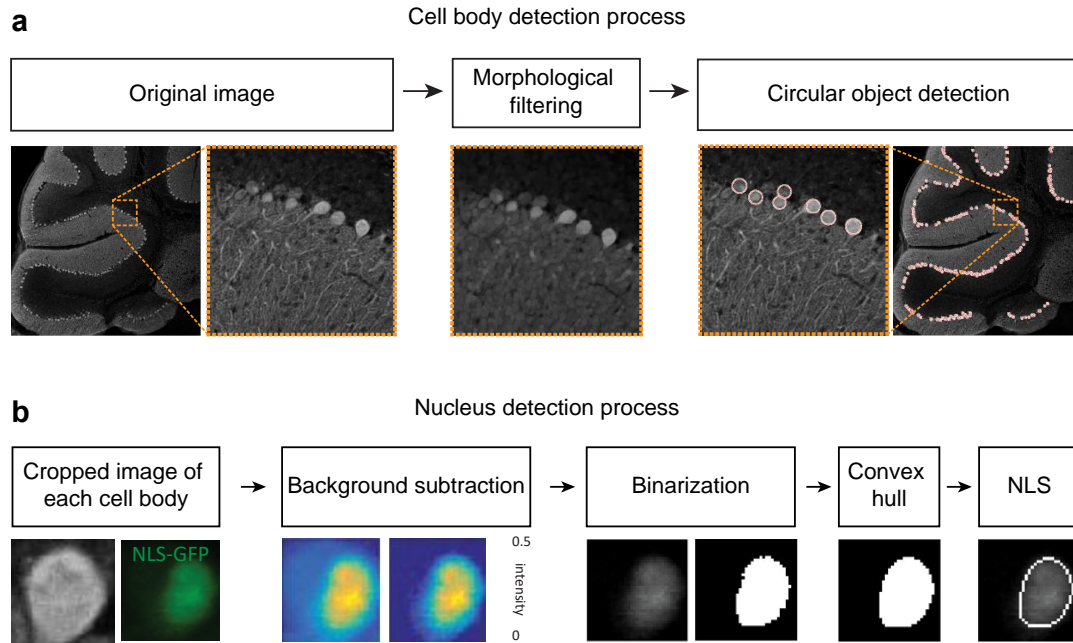
**Figure 5. AAV-PHP.eB can be used with gene regulatory elements to achieve cell type-restricted gene expression throughout the brain.** (a) Images show example cell type-specific promoters driving XFPs delivered as ssAAV-PHP.eB:Promoter-XFPs within the same animal (intravenous injection of  $1 \times 10^{12}$  vg per viral vector). Image is from the striatum of an adult C57Bl/6 mouse after 3 weeks of expression. (b) Ple67, a FEV/serotonin cell type-specific promoter driving GFP (ssAAV-PHP.eB:Ple67-GFP) was intravenously delivered at  $1 \times 10^{12}$  vg and colocalized to serotonergic neurons by IHC for serotonin (5-hydroxytryptamine, 5-HT, magenta). A whole midbrain coronal slice (left) is shown with the dorsal raphe nucleus (DRN) outlined in blue and expanded for detail (right). (c) A mouse tyrosine hydroxylase (TH) promoter (mTH) driving GFP (ssAAV-PHP.B:mTH-GFP) was intravenously injected at  $1 \times 10^{12}$  vg and imaging was performed after 2 weeks of expression. Native GFP fluorescence (green) is shown together with IHC for TH (magenta). Images show the substantia nigra pars compacta (SNc, left) and ventral tegmental area (VTA, right) in the midbrain. (d) A vector with a Purkinje cell-selective promoter (Ple155, Pcp2) driving GFP (ssAAV-PHP.eB:Ple155-GFP) was intravenously injected at  $1 \times 10^{12}$  vg and expression was examined at 4 weeks. Native GFP fluorescence (green) can be seen in Purkinje cells throughout the cerebellum within a whole sagittal section (left) and at higher resolution of the red boxed region (right). (e) A vector with a GABAergic interneuron specific enhancer within the forebrain driving nuclear-localized mRuby2 (mDlx-NLS-mRuby2) was intravenously injected at  $3 \times 10^{11}$  vg and expression was examined at 8 weeks. Native NLS-mRuby2 expression is seen within the forebrain (red, left) and colocalized to interneurons that are GAD67<sup>+</sup> in the cortex (green, right). The scale bars in (a; b, right; c; d, right; e, right) are 50  $\mu$ m. For (b and d, left) the scale bars are 1 mm. All panels are confocal images of native XFP fluorescence. (f) Summary table for FEV/Ple67, mTH, mDlx detailing the dosage, promoter length, transduction efficiency and specificity in the DRN, median raphe nucleus (MRN), VTA, SNc, cortex (CTX), and hippocampus (HIP).

## 4.6 Supplementary figures

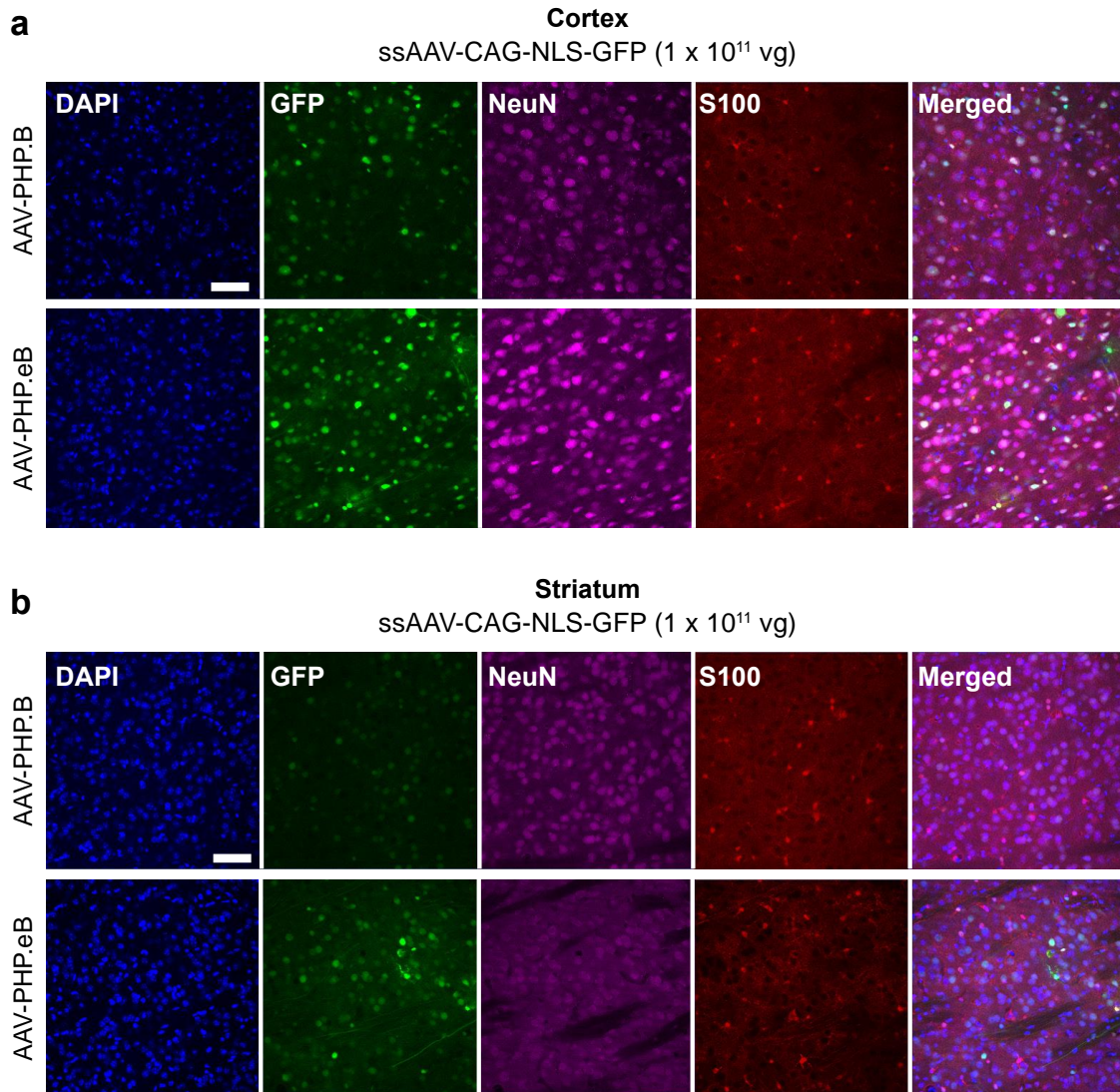


**Supplementary Figure 1. AAV-PHP.eB transduces several brain regions more efficiently than AAV-PHP.B.** (a-c) The cumulative distribution plots show native GFP

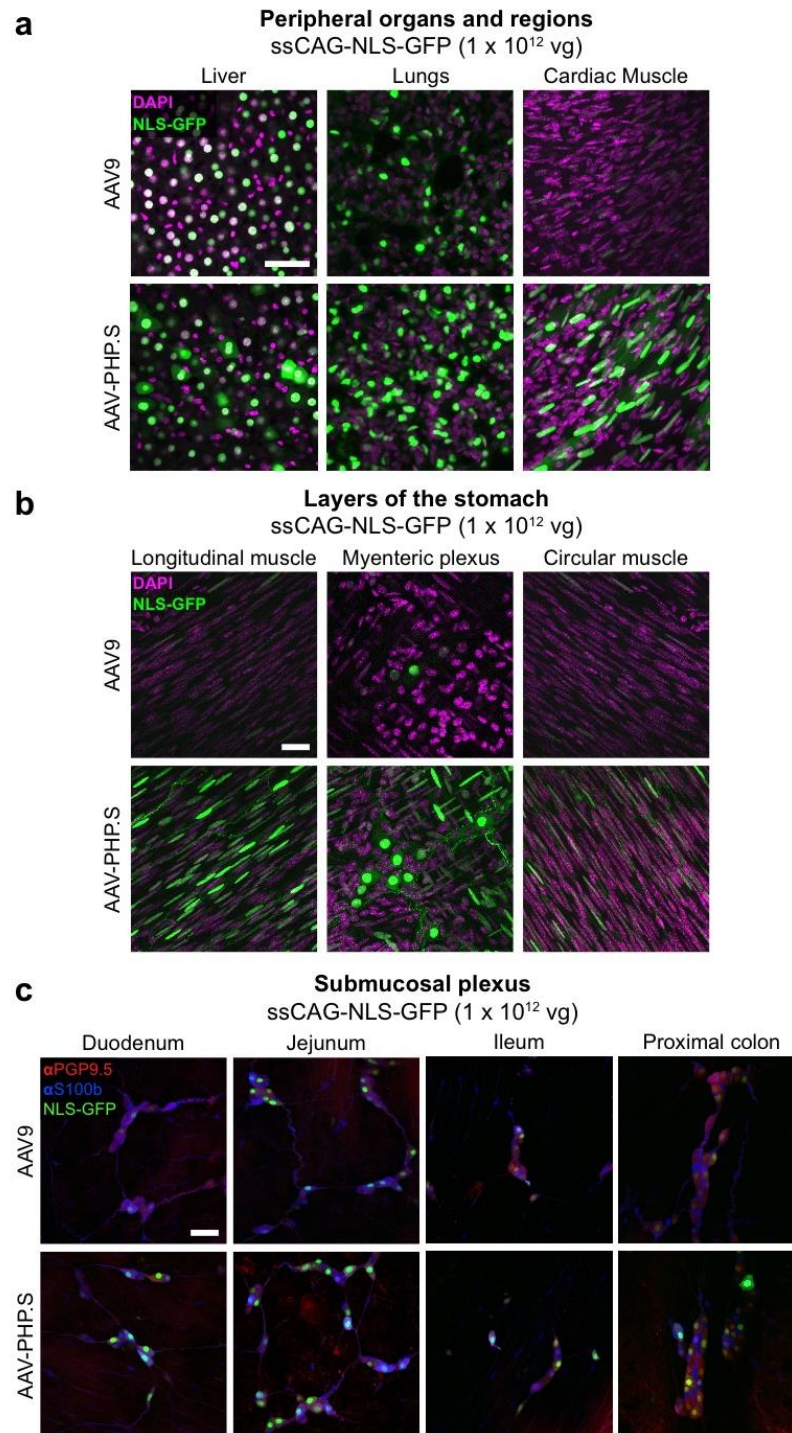
fluorescence, measured by the median NLS-GFP intensity within individual nuclei in the cortex (**a**), striatum (**b**) and within Calbindin<sup>+</sup> Purkinje cell bodies (**c**). In (**a** and **b**), ROIs were drawn manually around each DAPI<sup>+</sup> nuclei. In (**c**), the ROI detection was automated (see **Methods** and **Supplementary Fig. 2**). (**a-c**) The dotted lines represent the 50% intersect, where 50% of the cells have the corresponding median intensity or lower. (**c**) The solid black vertical line displays the cutoff point used during the automated detection to classify whether a cell had been transduced with ssAAV-CAG-NLS-GFP. Cells with a median intensity below or equal to 0.1 were classified as non-transduced while those above 0.1 were classified as transduced. (**a-c**) N = 4 individual animals per group.  $P \leq 0.0001$  for each region, Kolmogorov-Smirnov test.



**Supplementary Figure 2. An automated cell detection pipeline used for counting Purkinje cells in the cerebellum.** The pipeline used for automated cell body and nucleus detection. **(a)** To detect cell bodies from the original image, a morphological filter with a disk-shaped structural element is first used to blur thin fibrous processes (morphological filtering). A circular Hough transform was subsequently applied to the filtered image (circular object detection). Red circles indicate the objects detected each of which corresponds to a single cell body. **(b)** In **Fig. 2c**, we further analyzed the GFP channel to detect the nucleus of each cell. First, the same area at each cell body region was cropped from the GFP image (cropped image of each cell body). The background of each cropped image, estimated by 2D Gaussian smoothing with a standard deviation of 5-fold the cell body size, was subtracted (background subtraction). After thresholding the image to make it binary (binarization), the 2-D convex hull of the object (convex hull) was obtained as the final nuclear region (NLS).

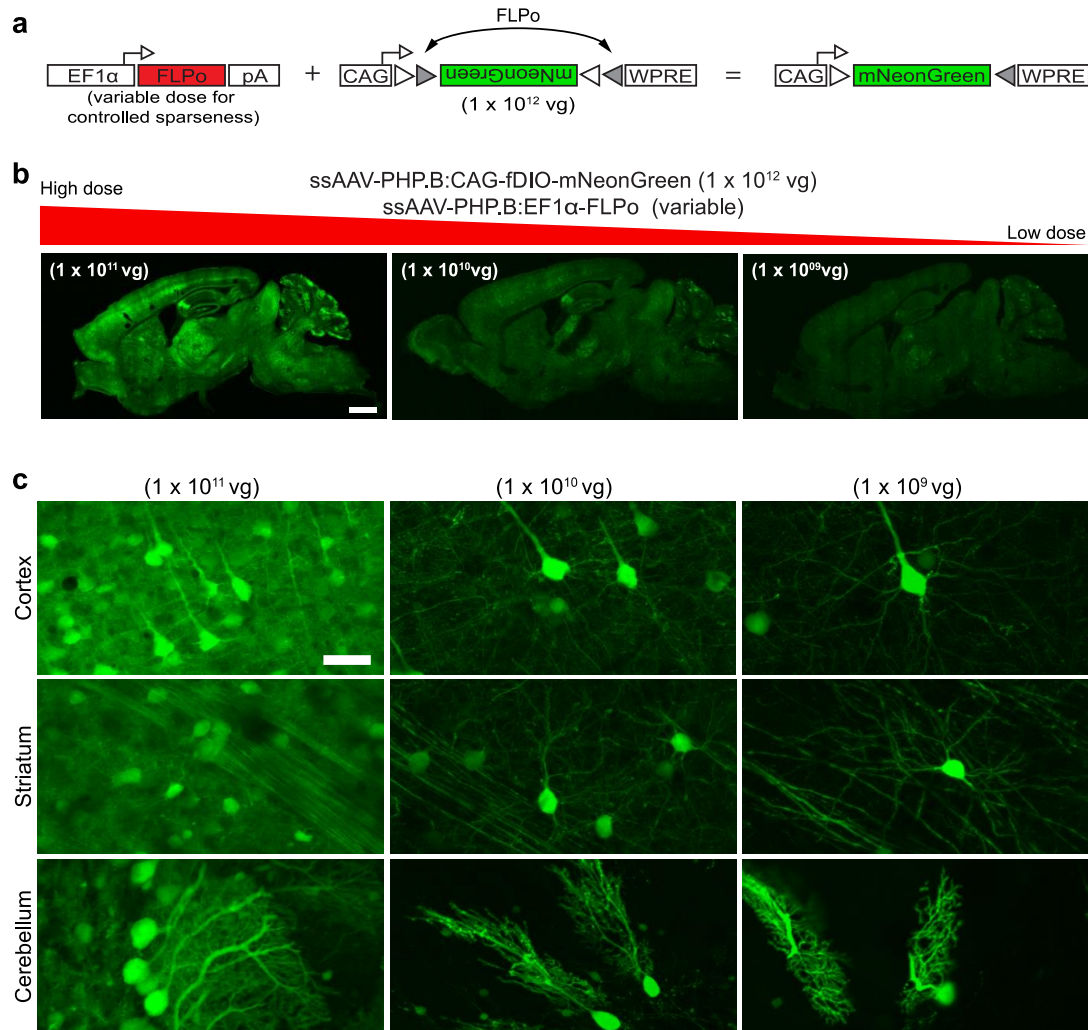


**Supplementary Figure 3. Representative immunohistochemistry in the cortex and striatum after transduction with AAV-PHP.B or AAV-PHP.eB.** Single-plane confocal images of native GFP fluorescence (green) after immunohistochemistry (IHC) with DAPI (DNA labeling, blue), NeuN (neuronal marker, magenta) and S100 (glial marker, red) in (a) cortex or (b) striatum. These data are presented in support of **Fig. 2**. All scale bars are 50  $\mu$ m. All imaging and display conditions for the GFP channel are matched across panel pairs to allow for comparisons. All images are from tissue extracted at three weeks post intravenous administration of ssAAV-CAG-NLS-GFP at  $1 \times 10^{11}$  vg to adult mice.

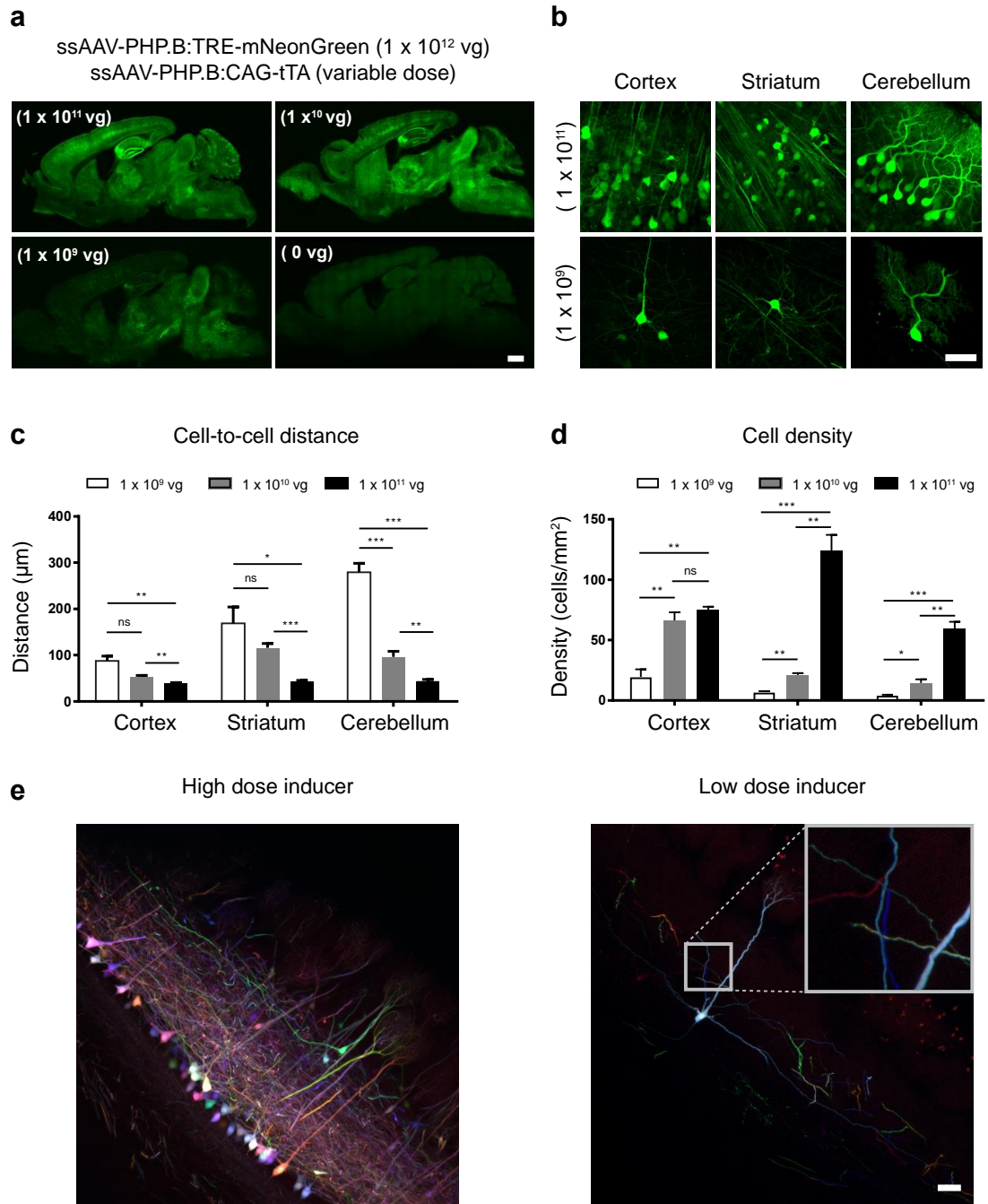


**Supplementary Figure 4. AAV-PHP.S transduces peripheral organs through intravenous delivery in adult mice. (a)** Representative images of the liver, lungs and

cardiac muscle transduced by AAV9 (top) or AAV-PHP.S (bottom). **(b)** Representative images of several layers of the stomach (longitudinal muscle, myenteric plexus, and circular muscle) transduced by AAV9 (top) or AAV-PHP.S (bottom). The myenteric plexus also contains muscle cells due to the uneven topography of the stomach. **(c)** Representative images of the submusocal plexus stained with neuronal marker PGP9.5 (red) and astrocyte marker S100b (blue) in the duodenum, jejunum, ileum, and proximal colon. **(a-c)** All images show native GFP fluorescence taken with confocal microscopy. Images in **(a)** are 20  $\mu\text{m}$  maximum intensity projections (MIPs) and images in **(b)** and **(c)** are single-plane images. All imaging and display conditions for the GFP channel are matched across panel pairs to allow for comparisons. All scale bars are 50  $\mu\text{m}$ . All images are from tissue extracted three weeks post intravenous administration of ssAAV-CAG-NLS-GFP at  $1 \times 10^{12}$  vg to adult mice. All experiments and imaging for direct comparisons were performed in parallel.

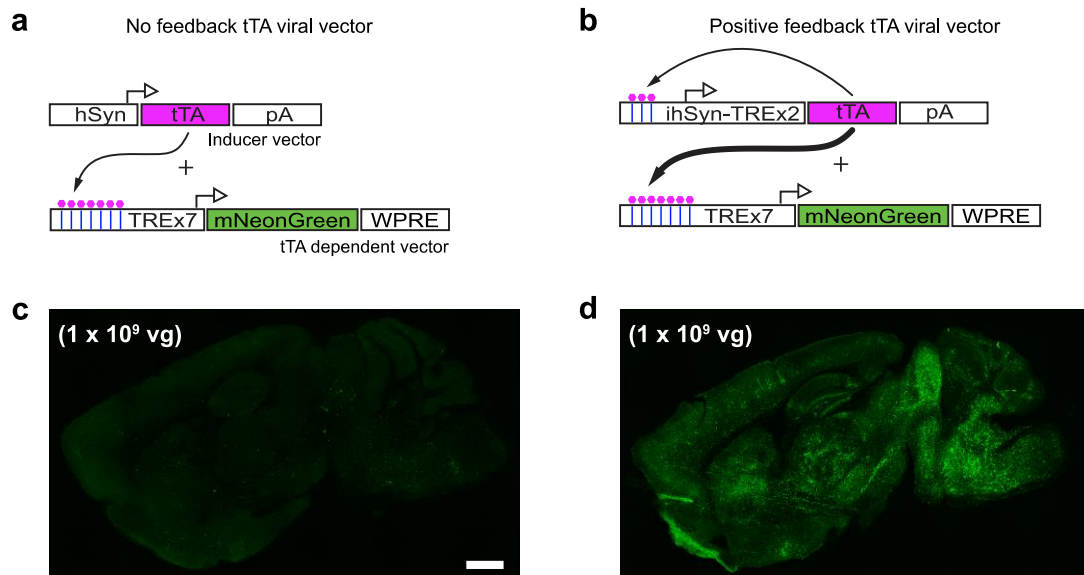


**Supplementary Figure 5. The FLPo-FRT two-component system allows for sparse gene expression.** (a) Schematic of the two viral vectors used for sparse gene expression with FLPo recombinase encoded in one viral vector and a transgene flanked by FRT sites on a second viral vector. In the basal state, the XFP is expressed in a double inverted orientation (fDIO) and upon the presence of FLPo is inverted into the orientation in line with the CAG promoter for expression. (b) Representative images of native mNeonGreen fluorescence two weeks after intravenous administration of ssAAV-PHP.B:CAG-fDIO-mNeonGreen at  $1 \times 10^{12}$  vg/mouse together with the indicated dose of ssAAV-PHP.B:EF1 $\alpha$ -FLPo. (c) Magnified images of the cortex, striatum, and cerebellum. Scale bars are 1 mm (b) and 50  $\mu$ m (c).



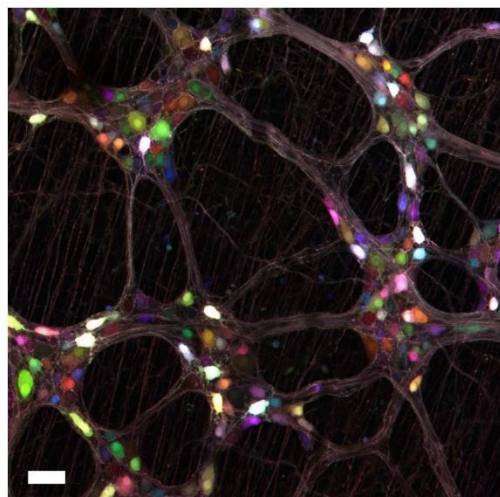
**Supplementary Figure 6. The use of the tetracycline-inducible system with viral vectors provides tunable dense-to-sparse gene expression.** (a) A two-component vector system consisting of one virus, ssAAV-PHP.B:TRE-mNeonGreen, delivered at  $1 \times 10^{12}$  vg into

adult mice, and a second virus, ssAAV-PHP.B:CAG-tTA, the dose of which was varied across mice. As a control, ssAAV-PHP.B:TRE-mNeonGreen was delivered alone to observe any basal expression from the ssAAV-PHP.B:TRE-mNeonGreen. All experimental, imaging and display conditions are matched across images. **(b)** Representative 20  $\mu\text{m}$  MIPs of native mNeonGreen fluorescence from the cortex, striatum, and cerebellum at high and low activation doses. **(c and d)** Measurements of the shortest cell-to-cell distance **(c)** and cell density **(d)** in the cortex, striatum and cerebellum.  $N = 3$  sections per group, mean  $\pm$  SEM, unpaired t-test ( $***p \leq 0.001$ ;  $**p \leq 0.01$ ;  $*p \leq 0.05$ ; n.s.,  $p \geq 0.05$ ). **(e)** A MIP image of a coronal section taken from the olfactory bulb of a Tbx21-Cre mouse transduced with a cocktail of ssAAV-PHP.eB:TRE-DIO-3XFP ( $1 \times 10^{12}$  vg total dose) and sAAV-PHP.eB:ihSyn1-tTA ( $1 \times 10^{10}$  vg – high dose) or ( $1 \times 10^8$  vg – low dose). In **(e, right)** the inset box in gray represents a high magnification of the region of interest outlined. Scale bars are 1 mm **(a)** and 50  $\mu\text{m}$  **(b and e)**. All images are native GFP fluorescence taken with confocal microscopy.

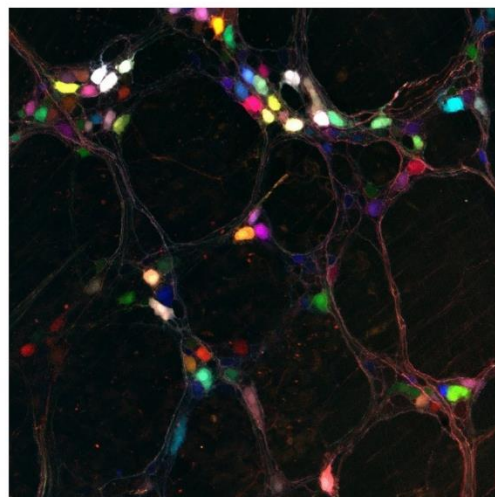


**Supplementary Figure 7. A positive feedback loop within the tTA inducer vector enhances gene expression.** Schematic of the two-component viral vector system without (a) and with (b) an inducible positive feedback loop with the viral vector containing the tTA. The schematic shows tTA protein (magenta hexagons) binding onto TRE elements (blue lines) to activate expression of (a) mNeonGreen or (b) mNeonGreen and the tTA containing vector. (c) Representative confocal images of sagittal brain sections transduced by the no-feedback vector and (d) positive-feedback tTA vector at  $1 \times 10^9$  vg. Both tTA inducers were co-administered with  $1 \times 10^{12}$  vg of the mNeonGreen viral vector. Scale bar in (c) is 1 mm.

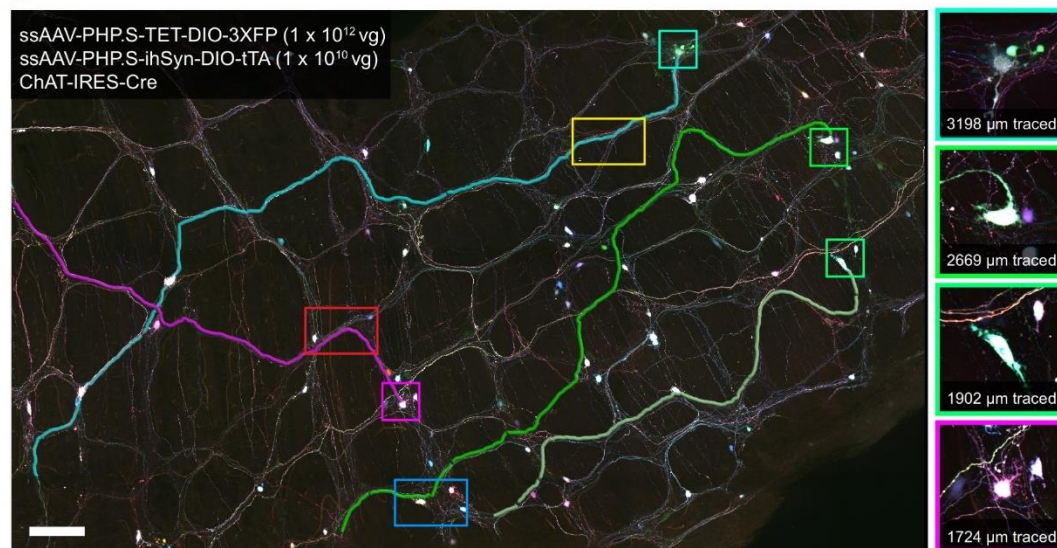
**a** Neuron-specific promoter  
ssAAV-PHP.S:hSyn1-3XFP



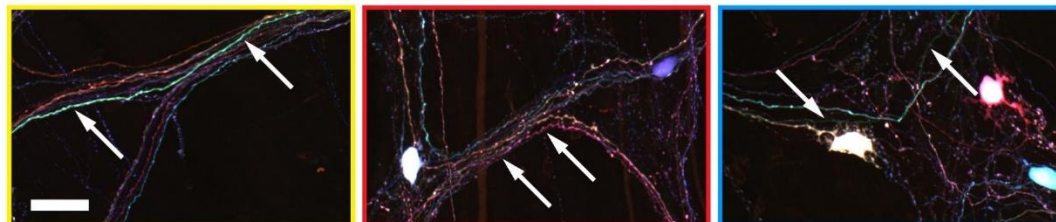
**b** ChAT-IRES-Cre transgenic  
ssAAV-PHP.S:CAG-DIO-3XFP



**c**



**d**



**Supplementary Figure 8. Cell type-restricted expression and sparse labeling in the enteric nervous system using AAV-PHP.S together with neuron-specific gene regulatory elements or Cre transgenic mice.** Images show XFP expression within the myenteric plexus. **(a)** ssAAV-PHP.S:hSyn1-3XFP was administered intravenously at a total dose of  $1 \times 10^{12}$  vg. **(b)** ssAAV-PHP.S:CAG-DIO-3XFP was administered intravenously at a total dose of  $1 \times 10^{12}$  vg to ChAT-IRES-Cre transgenic mice to achieve cell type-specific expression in cholinergic neurons. **(c)** Sparse labeling in the proximal colon of a ChAT-IRES-Cre transgenic using ssAAV-PHP.S-TET-DIO-3XFP at  $1 \times 10^{12}$  vg and ssAAV-PHP.S-ihSyn-DIO-tTA at  $1 \times 10^{10}$  vg. Color-coded square boxes display neurons traced manually in Imaris v8.3. **(d)** Rectangular boxes show higher magnification views of regions highlight how sparse multicolor labeling facilitates tracing of closely aligned processes. Double arrows mark traced processes. Scale bars in **(a and d)** are set to 50  $\mu\text{m}$  and in **(c)** to 200  $\mu\text{m}$ .

## 4.7 Supplementary tables

Supplementary Table 1 – Cre-selection lines, capsid sequences, and viral production efficiency					
	Cre selection line(s)	Amino acids*	Nucleotide sequence*	% of total recovered sequences	Average post purification yield per 150mm dish (vg)
<b>AAV9</b>	Parent	SAQ[7-mer]A	AGTGCCCAA__(7-mer insertion)__GCA		$1.28 \times 10^{12} \pm 5.97 \times 10^{11}$
<b>AAV-PHP.S</b>	GFAP-Cre	SAQ <b>QAVRTSLA</b>	AGTGCCCAAC <b>AGGCGGT</b> AGGAC <b>GTCTTT</b> GGCA	33%	$4.11 \times 10^{12} \pm 1.31 \times 10^{12}$
<b>AAV-PHP.B</b>	GFAP-Cre	SAQ <b>TLAVPFKA</b>	AGTGCCCAAC <b>CTTTGGCGGTGCCTTTTAAG</b> GCA		$2.45 \times 10^{12} \pm 1.02 \times 10^{12}$
<b>AAV-PHP.eB</b>	GFAP-, Vgat- and Vglut2-Cre	<b>SDGTLAVPFKA</b>	AGT <b>GATGGGACTTTGGCGGTGCCTTTTAAG</b> GCA	15.7%, 12.1%, and 11.1%, respectively	$3.06 \times 10^{12} \pm 1.18 \times 10^{12}$

\*Starting amino acid or codon for AAV9 = 586

**Supplementary Table 1. Capsid sequences, viral production efficiency, and Cre-selection lines.** The Cre-selection lines, amino acids, nucleotide sequences and average post-production yields for AAV9, AAV-PHP.S, AAV-PHP.B and AAV-PHP.eB.

	Applications within paper	Vector name pAAV-	Expression class
Tunable expression	<ul style="list-style-type: none"> <li>Expression of XFPs that are tunable in the density of expression in a constitutive, cell type-specific, or recombinase dependent manner (Fig. 4e,g), (Supplementary Fig. 5a-c), (Supplementary Fig. 6a,b), (Supplementary Fig. 7a-d), (Supplementary Fig. 8a-d).</li> </ul>	TRE <sup>a</sup> -mTurq2 TRE-mNeonG TRE-mRuby2	tTA-dependent
		TRE-DIO-mTurq2 TRE-DIO-mNeonG TRE-DIO-tdTomato TRE-DIO-mRuby2	Cre- and tTA-dependent
		CAG-fDIO-mNeonG	Flp dependent
		CAG <sup>b</sup> -tTA hSyn1 <sup>c</sup> -tTA ihSyn1 <sup>d</sup> -tTA ihSyn1-DIO-tTA	Constitutive
Tissue wide expression	<ul style="list-style-type: none"> <li>Expression of XFPs that are constitutive or Cre-dependent (Fig. 1c-g), (Fig. 2a-c), (Fig. 3a,c), (Fig. 4d), and (Supplementary Fig. 4a-c).</li> </ul>	CAG-mTurq2 CAG-mNeonG CAG-mRuby2 CAG-NLS-GFP-NLS	
		CAG-DIO-mTurq2 CAG-DIO-mNeonG CAG-DIO-mRuby2	Cre-dependent
Cell type-specific expression	<ul style="list-style-type: none"> <li>Expression of XFPs under the control of cell type-specific gene regulatory elements (Fig. 4a) and (Fig. 5a-e).</li> </ul>	hSyn1-mTurq2 hSyn1-mNeonG hSyn1-mRuby2 MBP <sup>e</sup> -eYFPf MsTH <sup>f</sup> -GFP GFAP <sup>g</sup> -mKate2.5f mDlx <sup>h</sup> -NLS-mRuby2	Cell type-specific
<p><sup>a</sup>TREpi = second generation tetracycline-regulated promoter</p> <p><sup>b</sup>CAG = synthetic promoter containing the cytomegalovirus early enhancer element, the promoter, first exon and first intron of chicken beta-actin gene, and the splice acceptor from the rabbit beta-globin gene</p> <p><sup>c</sup>hSyn1 = human Synapsin I promoter</p> <p><sup>d</sup>ihSyn1 = inducible intron human Synapsin I promoter</p> <p><sup>e</sup>MBP = myelin basic protein promoter</p> <p><sup>f</sup>MsTH = mouse tyrosine hydroxylase promoter</p> <p><sup>g</sup>GFAP = glial fibrillary acidic protein promoter</p> <p><sup>h</sup>mDlx = mouse distal-less homeobox promoter</p>			

**Supplementary Table 2. Summary of AAV resources.** A summary of the AAV plasmids used within the paper and their corresponding usages, with figure references.

#### 4.8 Supplementary movie caption

**Supplementary Movie 1. AAV-PHP.S transduction across multiple layers of the small intestine.** A video demonstrating the transduction of AAV-PHP.S across multiple layers of the small intestine after optical clearing with refractive index matching solution (RIMS). Native GFP fluorescence (green) from ssAAV-PHP.S:CAG-NLS-GFP at  $1 \times 10^{12}$ vg/mouse is shown. S100b staining (magenta) is shown for contrast. (doi: [10.22002/D1.237](https://doi.org/10.22002/D1.237))<sup>1</sup>

<sup>1</sup> = data file stored at [www.data.caltech.edu](http://www.data.caltech.edu)

## 4.9 Methods

### Animals

All procedures were approved by the California Institute of Technology Institutional Animal Care and Use Committee (IACUC). Mice were purchased from Jackson Laboratory (JAX) - C57Bl/6J mice (000664), GFAP-Cre (012886)<sup>114</sup>, Vgat-IRES-Cre (016962)<sup>115</sup>, Vglut2-IRES-Cre (016963)<sup>115</sup>, and Tbx21-Cre (024507)<sup>116</sup>. Male or female mice were used. Mice were randomly assigned to groups of predetermined sample size. All experiments with direct comparisons were performed in parallel to minimize variability.

### In vivo selection

The AAV-PHP.eB capsid was identified by selecting for AAV-PHP.B variants that transduced cells in Vglut2-IRES-Cre, Vgat-IRES-Cre, and GFAP-Cre mice through two rounds of selection as previously described<sup>18</sup>. Intravenous administration of AAV vectors was performed by injection into the retro-orbital sinus of adult mice (6-8 weeks of age). The AAV-PHP.B capsid was modified by partial randomization of the original AAV-PHP.B 7-mer insertion and 4 flanking AAs (AAV-PHP.B AA positions 587 - 597). The AAV-PHP.B-based library was generated by PCR using primer XF<sup>18</sup> and a mixture of 5 primers identical to the AAV-PHP.B template sequence with the exception of mixed bases at 3 consecutive randomized codons (NNK). In each primer mix, the position of the 3 randomized NNK codons was shifted by 6 bp (2 codons) such that each position in the 7-mer and the 2 AAs on each side of the 7-mer AAs were randomized, 3 AA at a time, in an overlapping manner. An equal amount of each of the 5 PCR products was assembled into the AAV-(delta)Cap-in-cis lox acceptor vector using the NEB Hi-Fi DNA Assembly Master Mix (New England Biolabs E2621) and used to generate the AAV-PHP.B-based AAV library. AAV-PHP.S was generated from an insertion of a random 7-mer between the 588-589 site of AAV9 and was identified from two rounds of *in vivo* selection in GFAP-Cre mice<sup>18</sup>.

## Plasmids

The backbones of all AAV plasmids that are double inverted orientation (DIO) were based on pAAV-Ef1a-DIO EYFP (Addgene, #27056), and all non-DIO plasmids were based on pAAV-EF1a-Cre (Addgene, #55636). The backbone of pAAV-CAG-fDIO-mNeonGreen was based on pAAV-Ef1a-fDIO EYFP (Addgene, #55641). The fluorescent protein cDNAs for mTurquoise2<sup>117</sup>, mNeonGreen<sup>118</sup>, mRuby2<sup>119</sup>, or mKate2.5 were synthesized as gBlocks (Integrated DNA Technologies). The MBP-eYFPf and GFAP-mkate2.5f vectors have a farnesylation sequence attached by overhang PCR. The pAAV-CAG-NLS-GFP vector was modified by PCR to add N- and C-terminal NLS sequences and assembled using NEB Hi-Fi DNA Assembly Master Mix (New England Biolabs E2621). The mouse tyrosine hydroxylase (TH) promoter (mTH) was based on the 2.5 kb region of a published rat TH promoter<sup>120</sup> and myelin basic promoter (MBP)<sup>121</sup>. Both were directly cloned from mouse genomic DNA. The hSyn1 promoter<sup>122</sup> was cloned from pAAV-hSyn Con/Foff EYFP (Addgene, #55651). The GFAP (GfABC<sub>1</sub>D) promoter was previously described<sup>123</sup> and was cloned upstream of a synthetic intron. The GFAP-mKate2.5 plasmid also contained 3 tandem copies of sequences complementary to 6 miRNAs: miR-1, miR-122, miR-375, miR-196a, miR-743 and miR-10b inserted between the WPRE and an SV40 pA sequence. The TRE was PCR amplified from a SG-TRE containing plasmid<sup>89</sup>. The inducible hSyn1 (ihSyn1) promoter was cloned by overlap PCR with the hSyn1 promoter, and oligonucleotides designed with a synthetic intron and 3 tetO binding sites. The AAV vectors expressing GFP from the Ple67 and Ple155 promoters were obtained from (Addgene, #49138 and #49140, respectively). The AAV-mDlx-NLS-mRuby vector was cloned by replacing the GFP reporter in pAAV-mDlx-GFP-Fishell-1 (Addgene, #83900) with a mRuby reporter fused with an N-terminal SV40 NLS.

## **Virus production and purification**

AAVs were generated in HEK 293T cells (ATCC) using Polyethylenimine (PEI)<sup>124</sup>. 72 hours post transfection, viral particles were harvested from the media and after 120 hours from cells and the media. Viral particles from the media were precipitated with 40% polyethylene glycol (Sigma, 89510-1KG-F) in 500 mM NaCl and combined with cell pellets for processing. The cell pellets were suspended in 500 mM NaCl, 40 mM Tris, 2.5 mM MgCl<sub>2</sub>, pH 8, and 100 U/mL of salt-activate nuclease (Arcticzymes) at 37°C for 30 minutes. Afterwards, the cells were clarified by centrifugation at 2,000 x g and then purified over iodixanol (Optiprep, Sigma; D1556) step gradients (15%, 25%, 40%, and 60%)<sup>125</sup>. Viruses were concentrated using Amicon filters (EMD, UFC910024), and formulated in sterile phosphate buffer saline (PBS). Virus titers were measured by determining the number of DNase I-resistant vg using qPCR using a linearized genome plasmid as a standard<sup>124</sup>.

## **Tissue preparation, immunohistochemistry and imaging**

Intravenous administration of AAV vectors was performed by injection into the retro-orbital sinus of adult mice (6-8 weeks of age). After allowing time for expression (3 weeks unless otherwise noted in Results), mice were anesthetized with EUTHASOL™ (Virbac) and transcardially perfused first with 0.1M phosphate buffer (PB, pH 7.4) at room temperature (RT) and then with freshly prepared, ice-cold 4% paraformaldehyde (PFA) in 0.1M PB. The tissues were post-fixed in 4% PFA overnight and afterwards transferred to phosphate buffered saline (PBS) with 0.05% sodium azide. The brains, lungs, and liver were sectioned to 100 µm, and the spinal cord to 300 µm, using a Leica VT1200 vibratome. Immunohistochemistry (IHC) was performed using 10% normal donkey serum (NDS), 0.1% Triton X-100 (no detergent was used with GAD67 staining) and 0.05% sodium azide. Primary antibodies used were rabbit Calbindin (1:500; Abcam, ab11426), mouse NeuN (1:500; EMD, MAB377), rabbit S100 (1:500; Abcam ab868), rabbit tyrosine hydroxylase (1:500; Abcam, ab112), goat 5-HT (1:200; Immunostar, 20079), rabbit PGP9.5 (1:500;

Abcam, ab10404), and mouse GAD67 (1:250; EMD, MAB5406). Sections were incubated with primary antibodies in blocking solution (PBS with 10% NDS and 0.05% sodium azide) for 16-24 hrs at RT and then washed in PBS overnight and followed by secondary antibody incubation (Jackson Immuno Research, 711-166-152, 711-606-152, 715-166-150, or 715-606-150) for 16-24 hrs at RT.

The dorsal root ganglia were post-fixed in 4% PFA at 4°C overnight and then cryopreserved in 30% sucrose at 4°C overnight. Tissues were then embedded in OCT (Tissue-Tek, Torrance, CA, USA), frozen in 2-methylbutane on dry ice and stored at -80°C until sectioning. For cryosectioning, the DRGs were cut at a thickness of 16 µm using a Leica CM3050 and adhered to a Superfrost Plus microscope slide (Fisher Scientific, Tustin, CA, USA). Slides were stored at -80°C until staining. For DRG IHC, the slides were incubated in rabbit anti-PGP9.5 antibody (1:250, Abcam, ab10404) in blocking solution (10% horse serum, 0.1% triton X-100, and 0.02% sodium azide in PBS) overnight at RT. The next day, sections were incubated with fluorescence-conjugated secondary antibody for 1 hr at RT (Jackson Immuno Research, 711-606-152). Between each step, sections were thoroughly washed with PBS. ProLong gold, anti-fade mounting medium (P36941; Molecular Probe, Life Technologies, Carlsbad, CA, USA) was applied to the slide before coverslip mounting.

The small intestines and colons were post-fixed in 4% PFA at 4°C overnight. Then the duodenum, jejunum, ileum, and proximal colon were sectioned and blocked in PBS with 0.1% triton X-100, 3% bovine serum albumin, and 0.01% sodium azide overnight at RT. The next day, the tissue was incubated with the mouse on mouse blocking reagent (Vector Laboratories, MKB-2213) in PBS with 0.01% sodium azide at RT for 16 hours. Afterwards, the tissue was stained in the initial blocking solution with guinea pig anti-PGP9.5 antibody (1:100, Abcam, ab10410) and rabbit anti-s100 beta antibody (1:200, Abcam, ab52642) for 48 hours and then washed in PBS with 0.01% sodium azide overnight. The secondary antibodies – goat anti-guinea pig Alexa 647 (A-21450) and goat anti-rabbit Alexa 555 (A-21428) Molecular Probes, Life Technologies, Carlsbad, CA, USA) – were incubated in the initial blocking solution for 48 hours. Unbound antibodies were then washed out with PBS

with 0.01% sodium azide overnight. Finally, samples were incubated in a refractive matching solution RIMS<sup>5</sup> solution for 6 hours before imaging.

For **Fig. 4b and g**, samples were made transparent and mounted using ScaleS4<sup>8</sup>. For **Fig. 4a,d, Fig. 5c, Supplementary Fig. 5a,b, and Fig. 6b,c** samples were made transparent and mounted using RIMS. All other samples were mounted with Prolong Diamond Antifade (ThermoFisher, P36965).

For some images, the 16-bit green channel (GFP) gamma was adjusted to enable visualization of low and high expressing cells while avoiding oversaturation. In all cases, changes to contrast or gamma as well as microscope laser settings remained constant across sets of images that were directly compared. Images were acquired on a Zeiss LSM 780 confocal microscope fitted with the following objectives: Fluar 5×/0.25 M27 Plan-Apochromat 10×/0.45 M27 (working distance 2.0 mm), and Plan-Apochromat 25×/0.8 Imm Corr DIC M27 multi-immersion or Zeiss Axiozoom V16 with a 0.7× objective fitted with a Hamamatsu c11440 camera.

### Image processing and data analysis

All image processing was performed using Zen Black 2012, Adobe Photoshop and Illustrator CC, ImageJ, neuTube Build 1.0z2016, and Bitplane Imaris 8.3, and custom-made scripts in Matlab. For data analysis, Microsoft Excel 2016, Matlab, Python, and GraphPad Prism 7.01 were used.

To mitigate any cross-talk between the 3 fluorescent proteins, linear unmixing was used (**Fig. 4d**). For a given imaging condition, the expected cross-talk between the three fluorophores was evaluated based on the emission and excitation spectra of each fluorophore. Then, for each pixel in the RGB image the corresponding unmixed signals were calculated using a non-negative least square fitting.

Cell profiling was done using Imaris (Bitplane). To detect neuronal cell bodies from the image data, we manually analyzed cells in the cortex and striatum by drawing a region of interest (ROI) around each cell's nucleus and intensity values were exported into

Microsoft Excel and Graphpad Prism for analysis. For analysis of the cerebellum, we developed a computational pipeline to facilitate the automation of the whole procedure, since the highly organized layout of Purkinje cells allowed for high accuracy automated counting (the true positive rate = 88.0% and the positive predictive value = 95.5% relative to manual detection)

To detect neuronal cell bodies from the image data (**Supplementary Fig. 2**), we applied a circular Hough transform to the maximum projection of the image stack. Prior to this process, morphological filtering with a circular structural element was performed with the images to remove thin fibrous processes. Each circular region detected from the images was considered as a single cell body. To measure the ratio of GFP-positive cells and their intensity in the cerebellum, we applied the abovementioned cell-body detection method to the Calbindin channel to detect the soma of Purkinje cells and used this to define ROIs in the GFP channel. We then subtracted the background from each ROI to exclude the expression (or leakage) of GFPs in cytosol. The background was estimated by applying a 2D Gaussian smoothing kernel to each cell body image. Subsequently, the image was converted to binary, and the convex hull of the connected object therein was calculated. The objects whose area were smaller than 10% of the cell body were removed, and we considered the largest object as a nucleus of interest if there were more than one object. Nuclei with a mean intensity higher than 0.1 (the intensity rescaled to the range of 0-1) was counted as GFP-positive. The mean intensity of the nucleus was calculated by averaging the pixel values inside the object (**Supplementary Fig. 1**).

The pipeline for automated cell body detection was also applied to other analyses (**Supplementary Fig. 7** and **Fig. 4f**). To measure the density of labeled cells in the sparse-labeling system (**Supplementary Fig. 6**), we applied the cell body detection method to the images and measured the density of detected cells by dividing the cell number by the area of the image. We also examined the distance between the neighboring cells; for each cell, all of the distances to other cells were calculated, and the minimum was selected as the distance to the closest neighbor. The values for all cells were then averaged for each region and condition. We used the same method to determine whether each cell expressed 1, 2, or 3

XPFs (**Fig. 4f**) in the one- and two-component labeling systems. To do this, after detecting each cell body, we collected all the normalized intensity values therein from each channel. If the mean intensity of the channel was above 0.1, we considered that this gene (color) was expressed in the cell. Counting the number of expressed channels in each cell, we calculated the fraction of cells according to the number of expressed gene types.

#### **4.10 Additional information**

#### **ACKNOWLEDGEMENTS**

We thank E. Mackey and K. Beadle for assistance with cloning and viral production, P. Anguiano for administrative assistance, and the entire Gradinaru group for discussions. We thank G. Stevens and V. Anand for their efforts in image analysis; P. Rajendra and S. Kalyanam at University of California Los Angeles and C. Fowlkes at the University of Irvine for useful discussions; the University of Pennsylvania vector core for the AAV2/9 Rep-Cap plasmid; M. Brenner at the University of Alabama for the GfABC<sub>1</sub>D promoter. pAAV-Ef1a-DIO EYFP, pAAV-EF1a-Cre and pAAV-Ef1a-fDIO EYFP were gifts from Karl Deisseroth (Addgene plasmid #27056, #55636 and #55641). pEMS2113 and pEMS2115 were gifts from Elizabeth Simpson (Addgene plasmid # 49138 and #49140).

This work was primarily supported by the National Institutes of Health (NIH) through grants to VG: Director's New Innovator 1DP2NS087949 and PECASE; SPARC 1OT2OD023848-01; NIH/National Institute on Aging (NIA) 1R01AG047664; NIH BRAIN 1U01NS090577; and NIH/ National Institute of Mental Health (NIMH) 1R21MH103824; and by a grant from the Hereditary Disease Foundation. Additional funding included the Gordon and Betty Moore Foundation through grant GBMF2809 to the Caltech Programmable Molecular Technology Initiative; the Curci Foundation; NIH U01 MH109147 02S1; NIH NS085910; and the Beckman Institute and Rosen Center at Caltech. S.K.M and V.G. are Heritage Principal Investigators supported by the Heritage Medical Research Institute.

All tools and protocols will be available through Addgene or the Beckman Institute for CLARITY, Optogenetics and Vector Engineering Research for technology development and broad dissemination:

<http://www.beckmaninstitute.caltech.edu/clover.shtml>.

## **CONTRIBUTIONS**

K.Y.C and B.E.D. designed and performed experiments, imaged samples, analyzed data. K.Y.C prepared figures with input from B.E.D and V.G. M.J.J. analyzed data, prepared figures, assisted with experiments and in manuscript preparation. B.B.Y. assisted with tissue processing, imaging, and virus production. A.G. helped with image analysis. N.R. assisted with molecular cloning. W.-L.W. assisted with tissue processing. L.S assisted in tissue processing. C.L and S.K.M assisted in experimental designs. K.Y.C, B.E.D and V.G. wrote the manuscript with support from all authors. B.E.D and V.G. conceived the project. V.G. supervised all aspects of the work.

## BIBLIOGRAPHY

- 1 Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science* **263**, 802-805 (1994).
- 2 Rajasethupathy, P., Ferenczi, E. & Deisseroth, K. Targeting Neural Circuits. *Cell* **165**, 524-534, doi:10.1016/j.cell.2016.03.047 (2016).
- 3 Spalteholz, W. *Über das durchsichtigmachen von menschlichen und tierischen präparaten und seine theoretischen bedingungen*. 2. aufl. edn, (1914).
- 4 Treweek, J. B. & Gradinaru, V. Extracting structural and functional features of widely distributed biological circuits with single cell resolution via tissue clearing and delivery vectors. *Curr Opin Biotechnol* **40**, 193-207, doi:10.1016/j.copbio.2016.03.012 (2016).
- 5 Chung, K. *et al.* Structural and molecular interrogation of intact biological systems. *Nature* **497**, 332-337, doi:10.1038/nature12107 (2013).
- 6 Yang, B. *et al.* Single-cell phenotyping within transparent intact tissue through whole-body clearing. *Cell* **158**, 945-958, doi:10.1016/j.cell.2014.07.017 (2014).
- 7 Pan, C. *et al.* Shrinkage-mediated imaging of entire organs and organisms using uDISCO. *Nat Methods* **13**, 859-867, doi:10.1038/nmeth.3964 (2016).
- 8 Susaki, E. A. *et al.* Whole-brain imaging with single-cell resolution using chemical cocktails and computational analysis. *Cell* **157**, 726-739, doi:10.1016/j.cell.2014.03.042 (2014).
- 9 Hama, H. *et al.* ScaleS: an optical clearing palette for biological imaging. *Nat Neurosci* **18**, 1518-1529, doi:10.1038/nn.4107 (2015).
- 10 Louveau, A. *et al.* Structural and functional features of central nervous system lymphatic vessels. *Nature* **523**, 337-341, doi:10.1038/nature14432 (2015).
- 11 Karsenty, G. & Ferron, M. The contribution of bone to whole-organism physiology. *Nature* **481**, 314-320, doi:10.1038/nature10763 (2012).
- 12 Mosialou, I. *et al.* MC4R-dependent suppression of appetite by bone-derived lipocalin 2. *Nature* **543**, 385-390, doi:10.1038/nature21697 (2017).
- 13 Goins, W. F., Huang, S., Cohen, J. B. & Glorioso, J. C. Engineering HSV-1 vectors for gene therapy. *Methods Mol Biol* **1144**, 63-79, doi:10.1007/978-1-4939-0428-0\_5 (2014).
- 14 Annoni, A., Goudy, K., Akbarpour, M., Naldini, L. & Roncarolo, M. G. Immune responses in liver-directed lentiviral gene therapy. *Transl Res* **161**, 230-240, doi:10.1016/j.trsl.2012.12.018 (2013).
- 15 Hastie, E. & Samulski, R. J. Adeno-associated virus at 50: a golden anniversary of discovery, research, and gene therapy success--a personal perspective. *Hum Gene Ther* **26**, 257-265, doi:10.1089/hum.2015.025 (2015).
- 16 Ferreira, V., Petry, H. & Salmon, F. Immune Responses to AAV-Vectors, the Glybera Example from Bench to Bedside. *Front Immunol* **5**, 82, doi:10.3389/fimmu.2014.00082 (2014).
- 17 Marchio, S., Sidman, R. L., Arap, W. & Pasqualini, R. Brain endothelial cell-targeted gene therapy of neurovascular disorders. *EMBO Mol Med* **8**, 592-594, doi:10.15252/emmm.201606407 (2016).

- 18 Choudhury, S. R. *et al.* Widespread Central Nervous System Gene Transfer and Silencing After Systemic Delivery of Novel AAV-AS Vector. *Mol Ther* **24**, 726-735, doi:10.1038/mt.2015.231 (2016).
- 19 Deverman, B. E. *et al.* Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain. *Nat Biotechnol* **34**, 204-209, doi:10.1038/nbt.3440 (2016).
- 20 Carabotti, M., Scirocco, A., Maselli, M. A. & Severi, C. The gut-brain axis: interactions between enteric microbiota, central and enteric nervous systems. *Ann Gastroenterol* **28**, 203-209 (2015).
- 21 Cai, D., Cohen, K. B., Luo, T., Lichtman, J. W. & Sanes, J. R. Improved tools for the Brainbow toolbox. *Nat Methods* **10**, 540-547 (2013).
- 22 Long, F. Building strong bones: molecular regulation of the osteoblast lineage. *Nat Rev Mol Cell Biol* **13**, 27-38, doi:10.1038/nrm3254 (2011).
- 23 Goltzman, D. Discoveries, drugs and skeletal disorders. *Nat Rev Drug Discov* **1**, 784-796, doi:10.1038/nrd916 (2002).
- 24 Bianco, P. & Robey, P. G. Skeletal stem cells. *Development* **142**, 1023-1027, doi:10.1242/dev.102210 (2015).
- 25 Klag, K. A. & Horton, W. A. Advances in treatment of achondroplasia and osteoarthritis. *Hum Mol Genet* **25**, R2-8, doi:10.1093/hmg/ddv419 (2016).
- 26 Morrison, S. J. & Scadden, D. T. The bone marrow niche for haematopoietic stem cells. *Nature* **505**, 327-334, doi:10.1038/nature12984 (2014).
- 27 Rachner, T. D., Khosla, S. & Hofbauer, L. C. Osteoporosis: now and the future. *Lancet* **377**, 1276-1287, doi:10.1016/S0140-6736(10)62349-5 (2011).
- 28 Beck, J. D. *et al.* Three-dimensional imaging of trabecular bone using the computer numerically controlled milling technique. *Bone* **21**, 281-287 (1997).
- 29 Slyfield, C. R., Tkachenko, E. V., Wilson, D. L. & Hernandez, C. J. Three-dimensional dynamic bone histomorphometry. *J Bone Miner Res* **27**, 486-495, doi:10.1002/jbmr.553 (2012).
- 30 Dodt, H. U. *et al.* Ultramicroscopy: three-dimensional visualization of neuronal networks in the whole mouse brain. *Nat Methods* **4**, 331-336, doi:10.1038/nmeth1036 (2007).
- 31 Erturk, A. *et al.* Three-dimensional imaging of solvent-cleared organs using 3DISCO. *Nat Protoc* **7**, 1983-1995, doi:10.1038/nprot.2012.119 (2012).
- 32 Becker, K., Jahrling, N., Saghaei, S., Weiler, R. & Dodt, H. U. Chemical clearing and dehydration of GFP expressing mouse brains. *PLoS One* **7**, e33916, doi:10.1371/journal.pone.0033916 (2012).
- 33 Richardson, D. S. & Lichtman, J. W. Clarifying Tissue Clearing. *Cell* **162**, 246-257, doi:10.1016/j.cell.2015.06.067 (2015).
- 34 Treweek, J. B. *et al.* Whole-body tissue stabilization and selective extractions via tissue-hydrogel hybrids for high-resolution intact circuit mapping and phenotyping. *Nat Protoc* **10**, 1860-1896, doi:10.1038/nprot.2015.122 (2015).
- 35 Susaki, E. A. & Ueda, H. R. Whole-body and Whole-Organ Clearing and Imaging Techniques with Single-Cell Resolution: Toward Organism-Level Systems Biology

- in Mammals. *Cell Chem Biol* **23**, 137-157, doi:10.1016/j.chembiol.2015.11.009 (2016).
- 36 Calve, S., Ready, A., Huppenbauer, C., Main, R. & Neu, C. P. Optical clearing in dense connective tissues to visualize cellular connectivity in situ. *PLoS One* **10**, e0116662, doi:10.1371/journal.pone.0116662 (2015).
- 37 Berke, I. M., Miola, J. P., David, M. A., Smith, M. K. & Price, C. Seeing through Musculoskeletal Tissues: Improving In Situ Imaging of Bone and the Lacunar Canalicular System through Optical Clearing. *PLoS One* **11**, e0150268, doi:10.1371/journal.pone.0150268 (2016).
- 38 Neu, C. P., Novak, T., Gilliland, K. F., Marshall, P. & Calve, S. Optical clearing in collagen- and proteoglycan-rich osteochondral tissues. *Osteoarthritis Cartilage* **23**, 405-413, doi:10.1016/j.joca.2014.11.021 (2015).
- 39 Genina, E. A., Bashkatov, A. N. & Tuchin, V. V. Optical clearing of human cranial bone by administration of immersion agents - art. no. 616311. *P Soc Photo-Opt Ins* **6163**, 16311-16311, doi:Artn 616311 10.1117/12.697308 (2006).
- 40 Acar, M. *et al.* Deep imaging of bone marrow shows non-dividing stem cells are mainly perisinusoidal. *Nature* **526**, 126-130, doi:10.1038/nature15250 (2015).
- 41 Susaki, E. A. *et al.* Advanced CUBIC protocols for whole-brain and whole-body clearing and imaging. *Nat Protoc* **10**, 1709-1727, doi:10.1038/nprot.2015.085 (2015).
- 42 Tainaka, K. *et al.* Whole-body imaging with single-cell resolution by tissue decolorization. *Cell* **159**, 911-924, doi:10.1016/j.cell.2014.10.034 (2014).
- 43 Tomer, R., Ye, L., Hsueh, B. & Deisseroth, K. Advanced CLARITY for rapid and high-resolution imaging of intact tissues. *Nat Protoc* **9**, 1682-1697, doi:10.1038/nprot.2014.123 (2014).
- 44 Artifacts of light. *Nat Meth* **10**, 1135-1135, doi:10.1038/nmeth.2760 (2013).
- 45 Reynaud, E. G., Peychl, J., Huisken, J. & Tomancak, P. Guide to light-sheet microscopy for adventurous biologists. *Nat Methods* **12**, 30-34, doi:10.1038/nmeth.3222 (2015).
- 46 Baumgart, E. & Kubitscheck, U. Scanned light sheet microscopy with confocal slit detection. *Opt Express* **20**, 21805-21814, doi:10.1364/OE.20.021805 (2012).
- 47 Soeda, T. *et al.* Sox9-expressing precursors are the cellular origin of the cruciate ligament of the knee joint and the limb tendons. *Genesis* **48**, 635-644, doi:10.1002/dvg.20667 (2010).
- 48 Elefteriou, F. & Yang, X. Genetic mouse models for bone studies--strengths and limitations. *Bone* **49**, 1242-1254, doi:10.1016/j.bone.2011.08.021 (2011).
- 49 Ono, N., Ono, W., Nagasawa, T. & Kronenberg, H. M. A subset of chondrogenic cells provides early mesenchymal progenitors in growing bones. *Nat Cell Biol* **16**, 1157-1167, doi:10.1038/ncb3067 (2014).
- 50 Muhlfield, C., Hegermann, J., Wrede, C. & Ochs, M. A review of recent developments and applications of morphometry/stereology in lung research. *Am J Physiol Lung Cell Mol Physiol* **309**, L526-536, doi:10.1152/ajplung.00047.2015 (2015).

- 51 Gundersen, H. J. & Jensen, E. B. The efficiency of systematic sampling in stereology and its prediction. *J Microsc* **147**, 229-263 (1987).
- 52 Gundersen, H. J., Jensen, E. B., Kieu, K. & Nielsen, J. The efficiency of systematic sampling in stereology--reconsidered. *J Microsc* **193**, 199-211 (1999).
- 53 Boyce, R. W., Dorph-Petersen, K. A., Lyck, L. & Gundersen, H. J. Design-based stereology: introduction to basic concepts and practical approaches for estimation of cell number. *Toxicol Pathol* **38**, 1011-1025, doi:10.1177/0192623310385140 (2010).
- 54 Sterio, D. C. The unbiased estimation of number and sizes of arbitrary particles using the disector. *J Microsc* **134**, 127-136 (1984).
- 55 Gundersen, H. J. Stereology of arbitrary particles. A review of unbiased number and size estimators and the presentation of some new ones, in memory of William R. Thompson. *J Microsc* **143**, 3-45 (1986).
- 56 Gundersen, H. J. *et al.* The new stereological tools: disector, fractionator, nucleator and point sampled intercepts and their use in pathological research and diagnosis. *APMIS* **96**, 857-881 (1988).
- 57 Gardi, J. E., Nyengaard, J. R. & Gundersen, H. J. The proportionator: unbiased stereological estimation using biased automatic image analysis and non-uniform probability proportional to size sampling. *Comput Biol Med* **38**, 313-328, doi:10.1016/j.combiomed.2007.11.002 (2008).
- 58 Kawai, M., Modder, U. I., Khosla, S. & Rosen, C. J. Emerging therapeutic opportunities for skeletal restoration. *Nat Rev Drug Discov* **10**, 141-156, doi:10.1038/nrd3299 (2011).
- 59 Taylor, S. *et al.* Time-dependent cellular and transcriptional changes in the osteoblast lineage associated with sclerostin antibody treatment in ovariectomized rats. *Bone* **84**, 148-159, doi:10.1016/j.bone.2015.12.013 (2016).
- 60 Nioi, P. *et al.* Transcriptional Profiling of Laser Capture Microdissected Subpopulations of the Osteoblast Lineage Provides Insight Into the Early Response to Sclerostin Antibody in Rats. *J Bone Miner Res* **30**, 1457-1467, doi:10.1002/jbmr.2482 (2015).
- 61 Dempster, D. W. Osteoporosis and the burden of osteoporosis-related fractures. *Am J Manag Care* **17 Suppl 6**, S164-169 (2011).
- 62 Shah, S. *et al.* Single-molecule RNA detection at depth by hybridization chain reaction and tissue hydrogel embedding and clearing. *Development* **143**, 2862-2867, doi:10.1242/dev.138560 (2016).
- 63 Shaner, N. C., Steinbach, P. A. & Tsien, R. Y. A guide to choosing fluorescent proteins. *Nat Methods* **2**, 905-909, doi:10.1038/nmeth819 (2005).
- 64 Aspelund, A. *et al.* A dural lymphatic vascular system that drains brain interstitial fluid and macromolecules. *J Exp Med* **212**, 991-999, doi:10.1084/jem.20142290 (2015).
- 65 Suzuki, K. *et al.* In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* **540**, 144-149, doi:10.1038/nature20565 (2016).

- 66 Yang, Y. *et al.* A dual AAV system enables the Cas9-mediated correction of a metabolic liver disease in newborn mice. *Nat Biotechnol* **34**, 334-338, doi:10.1038/nbt.3469 (2016).
- 67 Urban, D. J. & Roth, B. L. DREADDs (designer receptors exclusively activated by designer drugs): chemogenetic tools with therapeutic utility. *Annu Rev Pharmacol Toxicol* **55**, 399-417, doi:10.1146/annurev-pharmtox-010814-124803 (2015).
- 68 Marshall, J. D. *et al.* Cell-Type-Specific Optical Recording of Membrane Voltage Dynamics in Freely Moving Mice. *Cell* **167**, 1650-1662 e1615, doi:10.1016/j.cell.2016.11.021 (2016).
- 69 Kim, C. K. *et al.* Simultaneous fast measurement of circuit dynamics at multiple sites across the mammalian brain. *Nat Methods* **13**, 325-328, doi:10.1038/nmeth.3770 (2016).
- 70 Van der Perren, A., Van den Haute, C. & Baekelandt, V. Viral vector-based models of Parkinson's disease. *Curr Top Behav Neurosci* **22**, 271-301, doi:10.1007/7854\_2014\_310 (2015).
- 71 Hocquemiller, M., Giersch, L., Audrain, M., Parker, S. & Cartier, N. Adeno-Associated Virus-Based Gene Therapy for CNS Diseases. *Hum Gene Ther* **27**, 478-496, doi:10.1089/hum.2016.087 (2016).
- 72 Samulski, R. J. & Muzyczka, N. AAV-Mediated Gene Therapy for Research and Therapeutic Purposes. *Annu Rev Virol* **1**, 427-451, doi:10.1146/annurev-virology-031413-085355 (2014).
- 73 Bennett, J. *et al.* Safety and durability of effect of contralateral-eye administration of AAV2 gene therapy in patients with childhood-onset blindness caused by RPE65 mutations: a follow-on phase 1 trial. *Lancet* **388**, 661-672, doi:10.1016/S0140-6736(16)30371-3 (2016).
- 74 Iyer, S. M. *et al.* Virally mediated optogenetic excitation and inhibition of pain in freely moving nontransgenic mice. *Nat Biotechnol* **32**, 274-278, doi:10.1038/nbt.2834 (2014).
- 75 Samad, O. A. *et al.* Virus-mediated shRNA knockdown of Na(v)1.3 in rat dorsal root ganglion attenuates nerve injury-induced neuropathic pain. *Mol Ther* **21**, 49-56, doi:10.1038/mt.2012.169 (2013).
- 76 Fischer, G., Pan, B., Vilceanu, D., Hogan, Q. H. & Yu, H. Sustained relief of neuropathic pain by AAV-targeted expression of CBD3 peptide in rat dorsal root ganglion. *Gene Ther* **21**, 44-51, doi:10.1038/gt.2013.56 (2014).
- 77 Chang, R. B., Strohlic, D. E., Williams, E. K., Umans, B. D. & Liberles, S. D. Vagal Sensory Neuron Subtypes that Differentially Control Breathing. *Cell* **161**, 622-633, doi:10.1016/j.cell.2015.03.022 (2015).
- 78 Williams, E. K. *et al.* Sensory Neurons that Detect Stretch and Nutrients in the Digestive System. *Cell* **166**, 209-221, doi:10.1016/j.cell.2016.05.011 (2016).
- 79 Eldridge, M. A. *et al.* Chemogenetic disconnection of monkey orbitofrontal and rhinal cortex reversibly disrupts reward value. *Nat Neurosci* **19**, 37-39, doi:10.1038/nn.4192 (2016).
- 80 Foust, K. D. *et al.* Intravascular AAV9 preferentially targets neonatal neurons and adult astrocytes. *Nat Biotechnol* **27**, 59-65, doi:10.1038/nbt.1515 (2009).

- 81 Yang, B. *et al.* Global CNS transduction of adult mice by intravenously delivered rAAVrh.8 and rAAVrh.10 and nonhuman primates by rAAVrh.10. *Mol Ther* **22**, 1299-1309, doi:10.1038/mt.2014.68 (2014).
- 82 Pulicherla, N. *et al.* Engineering liver-detargeted AAV9 vectors for cardiac and musculoskeletal gene transfer. *Mol Ther* **19**, 1070-1078, doi:10.1038/mt.2011.22 (2011).
- 83 Betley, J. N. & Sternson, S. M. Adeno-associated viral vectors for mapping, monitoring, and manipulating neural circuits. *Hum Gene Ther* **22**, 669-677, doi:10.1089/hum.2010.204 (2011).
- 84 Yonehara, K. *et al.* Congenital Nystagmus Gene FRMD7 Is Necessary for Establishing a Neuronal Circuit Asymmetry for Direction Selectivity. *Neuron* **89**, 177-193, doi:10.1016/j.neuron.2015.11.032 (2016).
- 85 Beier, K. T. *et al.* Circuit Architecture of VTA Dopamine Neurons Revealed by Systematic Input-Output Mapping. *Cell* **162**, 622-634, doi:10.1016/j.cell.2015.07.015 (2015).
- 86 Duque, S. *et al.* Intravenous administration of self-complementary AAV9 enables transgene delivery to adult motor neurons. *Mol Ther* **17**, 1187-1196, doi:10.1038/mt.2009.71 (2009).
- 87 Gombash, S. E. *et al.* Intravenous AAV9 efficiently transduces myenteric neurons in neonate and juvenile mice. *Front Mol Neurosci* **7**, 81, doi:10.3389/fnmol.2014.00081 (2014).
- 88 Inagaki, K. *et al.* Robust systemic transduction with AAV9 vectors in mice: efficient global cardiac gene transfer superior to that of AAV8. *Mol Ther* **14**, 45-53, doi:10.1016/j.ymthe.2006.03.014 (2006).
- 89 Parekh, R. & Ascoli, G. A. Neuronal morphology goes digital: a research hub for cellular and system neuroscience. *Neuron* **77**, 1017-1038, doi:10.1016/j.neuron.2013.03.008 (2013).
- 90 Agha-Mohammadi, S. *et al.* Second-generation tetracycline-regulatable promoter: repositioned tet operator elements optimize transactivator synergy while shorter minimal promoter offers tight basal leakiness. *J Gene Med* **6**, 817-828, doi:10.1002/jgm.566 (2004).
- 91 Raymond, C. S. & Soriano, P. High-efficiency FLP and PhiC31 site-specific recombination in mammalian cells. *PLoS One* **2**, e162, doi:10.1371/journal.pone.0000162 (2007).
- 92 Gallia, G. L. & Khalili, K. Evaluation of an autoregulatory tetracycline regulated system. *Oncogene* **16**, 1879-1884, doi:10.1038/sj.onc.1201706 (1998).
- 93 de Leeuw, C. N. *et al.* rAAV-compatible MiniPromoters for restricted expression in the brain and eye. *Mol Brain* **9**, 52, doi:10.1186/s13041-016-0232-4 (2016).
- 94 Dimidschstein, J. *et al.* A viral strategy for targeting and manipulating interneurons across vertebrate species. *Nat Neurosci*, doi:10.1038/nn.4430 (2016).
- 95 Towne, C., Pertin, M., Beggah, A. T., Aebischer, P. & Decosterd, I. Recombinant adeno-associated virus serotype 6 (rAAV2/6)-mediated gene transfer to nociceptive neurons through different routes of delivery. *Mol Pain* **5**, 52, doi:10.1186/1744-8069-5-52 (2009).

- 96 Fischer, G. *et al.* Direct injection into the dorsal root ganglion: technical, behavioral, and histological observations. *J Neurosci Methods* **199**, 43-55, doi:10.1016/j.jneumeth.2011.04.021 (2011).
- 97 Tang, J. C. *et al.* Cell type-specific manipulation with GFP-dependent Cre recombinase. *Nat Neurosci* **18**, 1334-1341, doi:10.1038/nn.4081 (2015).
- 98 Mingozzi, F. & High, K. A. Immune responses to AAV vectors: overcoming barriers to successful gene therapy. *Blood* **122**, 23-36, doi:10.1182/blood-2013-01-306647 (2013).
- 99 Manno, C. S. *et al.* Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nat Med* **12**, 342-347, doi:10.1038/nm1358 (2006).
- 100 Tervo, D. G. *et al.* A Designer AAV Variant Permits Efficient Retrograde Access to Projection Neurons. *Neuron* **92**, 372-382, doi:10.1016/j.neuron.2016.09.021 (2016).
- 101 Nathanson, J. L. *et al.* Short Promoters in Viral Vectors Drive Selective Expression in Mammalian Inhibitory Neurons, but do not Restrict Activity to Specific Inhibitory Cell-Types. *Front Neural Circuits* **3**, 19, doi:10.3389/neuro.04.019.2009 (2009).
- 102 Jackson, K. L., Dayton, R. D., Deverman, B. E. & Klein, R. L. Better Targeting, Better Efficiency for Wide-Scale Neuronal Transduction with the Synapsin Promoter and AAV-PHP.B. *Front Mol Neurosci* **9**, 116, doi:10.3389/fnmol.2016.00116 (2016).
- 103 Micheva, K. D., Busse, B., Weiler, N. C., O'Rourke, N. & Smith, S. J. Single-synapse analysis of a diverse synapse population: proteomic imaging methods and markers. *Neuron* **68**, 639-653, doi:10.1016/j.neuron.2010.09.024 (2010).
- 104 Mikuni, T., Nishiyama, J., Sun, Y., Kamasawa, N. & Yasuda, R. High-Throughput, High-Resolution Mapping of Protein Localization in Mammalian Brain by In Vivo Genome Editing. *Cell* **165**, 1803-1817, doi:10.1016/j.cell.2016.04.044 (2016).
- 105 Marshel, J. H., Mori, T., Nielsen, K. J. & Callaway, E. M. Targeting single neuronal networks for gene expression and cell labeling in vivo. *Neuron* **67**, 562-574, doi:10.1016/j.neuron.2010.08.001 (2010).
- 106 Wertz, A. *et al.* PRESYNAPTIC NETWORKS. Single-cell-initiated monosynaptic tracing reveals layer-specific cortical network modules. *Science* **349**, 70-74, doi:10.1126/science.aab1687 (2015).
- 107 Schwarz, L. A. *et al.* Viral-genetic tracing of the input-output organization of a central noradrenaline circuit. *Nature* **524**, 88-92, doi:10.1038/nature14600 (2015).
- 108 Loulier, K. *et al.* Multiplex cell and lineage tracking with combinatorial labels. *Neuron* **81**, 505-520, doi:10.1016/j.neuron.2013.12.016 (2014).
- 109 Kitamura, K., Judkewitz, B., Kano, M., Denk, W. & Hausser, M. Targeted patch-clamp recordings and single-cell electroporation of unlabeled neurons in vivo. *Nat Methods* **5**, 61-67, doi:10.1038/nmeth1150 (2008).
- 110 Haas, K., Sin, W. C., Javaherian, A., Li, Z. & Cline, H. T. Single-cell electroporation for gene transfer in vivo. *Neuron* **29**, 583-591 (2001).
- 111 Young, P. *et al.* Single-neuron labeling with inducible Cre-mediated knockout in transgenic mice. *Nat Neurosci* **11**, 721-728, doi:10.1038/nn.2118 (2008).

- 112 Badea, T. C. *et al.* New mouse lines for the analysis of neuronal morphology using CreER(T)/loxP-directed sparse labeling. *PLoS One* **4**, e7859, doi:10.1371/journal.pone.0007859 (2009).
- 113 Lin, M. Z. & Schnitzer, M. J. Genetically encoded indicators of neuronal activity. *Nat Neurosci* **19**, 1142-1153, doi:10.1038/nn.4359 (2016).
- 114 Sofroniew, N. J., Flickinger, D., King, J. & Svoboda, K. A large field of view two-photon mesoscope with subcellular resolution for in vivo imaging. *Elife* **5**, doi:10.7554/eLife.14472 (2016).
- 115 Garcia, A. D., Doan, N. B., Imura, T., Bush, T. G. & Sofroniew, M. V. GFAP-expressing progenitors are the principal source of constitutive neurogenesis in adult mouse forebrain. *Nat Neurosci* **7**, 1233-1241, doi:10.1038/nn1340 (2004).
- 116 Vong, L. *et al.* Leptin action on GABAergic neurons prevents obesity and reduces inhibitory tone to POMC neurons. *Neuron* **71**, 142-154, doi:10.1016/j.neuron.2011.05.028 (2011).
- 117 Haddad, R. *et al.* Olfactory cortical neurons read out a relative time code in the olfactory bulb. *Nat Neurosci* **16**, 949-957, doi:10.1038/nn.3407 (2013).
- 118 Goedhart, J. *et al.* Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. *Nat Commun* **3**, 751, doi:10.1038/ncomms1738 (2012).
- 119 Shaner, N. C. *et al.* A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum*. *Nat Methods* **10**, 407-409, doi:10.1038/nmeth.2413 (2013).
- 120 Lam, A. J. *et al.* Improving FRET dynamic range with bright green and red fluorescent proteins. *Nat Methods* **9**, 1005-1012, doi:10.1038/nmeth.2171 (2012).
- 121 Oh, M. S., Hong, S. J., Huh, Y. & Kim, K. S. Expression of transgenes in midbrain dopamine neurons using the tyrosine hydroxylase promoter. *Gene Ther* **16**, 437-440, doi:10.1038/gt.2008.148 (2009).
- 122 Gow, A., Friedrich, V. L., Jr. & Lazzarini, R. A. Myelin basic protein gene contains separate enhancers for oligodendrocyte and Schwann cell expression. *J Cell Biol* **119**, 605-616 (1992).
- 123 Kugler, S., Kilic, E. & Bahr, M. Human synapsin 1 gene promoter confers highly neuron-specific long-term transgene expression from an adenoviral vector in the adult rat brain depending on the transduced area. *Gene Ther* **10**, 337-347, doi:10.1038/sj.gt.3301905 (2003).
- 124 Lee, Y., Messing, A., Su, M. & Brenner, M. GFAP promoter elements required for region-specific and astrocyte-specific expression. *Glia* **56**, 481-493, doi:10.1002/glia.20622 (2008).
- 125 Gray, S. J. *et al.* Production of recombinant adeno-associated viral vectors and use in in vitro and in vivo administration. *Curr Protoc Neurosci* **Chapter 4**, Unit 4 17, doi:10.1002/0471142301.ns0417s57 (2011).
- 126 Zolotukhin, S. *et al.* Recombinant adeno-associated virus purification using novel methods improves infectious titer and yield. *Gene Ther* **6**, 973-985, doi:10.1038/sj.gt.3300938 (1999).